DOCUMENT RESUME

ED 225 831 SE 040 190

AUTHOR Dougherty, Knowles; Herbert, Martin
TITLE Preliminary Test Data: Fifth Grade. Evaluation Report 6-B-2. Extended Pilot Trial of the Comprehensive School Mathematics Program.
INSTITUTION CEMREL, Inc., St. Louis, Mo.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
PUB DATE Oct 79
NOTE 89p.; For related documents, see SE 040 181-196 and SE 040 340-348.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC04 Plus Postage.
DESCRIPTORS Achievement; Educational Research; *Elementary School Mathematics; Evaluation Methods; *Grade 5; Intermediate Grades; *Mathematics; Mathematics Curriculum; Mathematics Instruction; *Program Evaluation; Testing
IDENTIFIERS *Comprehensive School Mathematics Program; *Mathematics Education Research

ABSTRACT

The Comprehensive School Mathematics Program (CSMP) is a program of CEMREL, Inc., one of the national educational laboratories, and was funded by the National Institute of Education (NIE). Its major purpose is the development of curriculum materials for grades kindergarten through 6. This is a report of the first year of a 2-year evaluation of the CSMP fifth-grade materials. The study was intended as both a pilot test of certain scales and a preliminary comparative evaluation of CSMP/non-CSMP achievement. A small group of St. Louis area classes were chosen for the pilot study. Main results noted: 1) Significant differences in favor of CSMP classes in mental arithmetic, number relations, number systems (other than whole numbers), certain word problems, and problems with multiple correct answers; 2) no significant differences on standard computation, geometry, probability, and place value; 3) similar results for comparisons of the two groups by sex and at different levels of reading ability; 4) a consistency with previous findings, particularly those in large-scale fourth-grade comparisons. (MP)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

LIMITED FIELD TRIALS OF THE COMPREHENSIVE SCHOOL MATHEMATICS PROGRAM:

EVALUATION REPORT SERIES

Evaluation Report 6-B-2
Preliminary Test Data: Fifth Grade

ED225831

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

✓ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
CEMREL INC.
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

SE 040 190

Extended Pilot Trial of the Comprehensive School Mathematics Program

Evaluation Report 6-B-2
Preliminary Test Data: Fifth Grade

Knowles Dougherty
Martin Hebert

October, 1979

Developed by CEMREL, Inc., a private nonprofit corporation supported in part as an educational laboratory by funds from the National Institute of Education, Department of Health, Education, and Welfare. The opinions expressed in this publication do not necessarily reflect the position or policy of the National Institute of Education, and no official endorsement should be inferred.

Copyright on these materials is claimed only during the period of development, test, and evaluation, unless additional authorization is granted by the National Institute of Education, to claim copyright on the final materials. For information on the status of the copyright claim, contact either the copyright proprietor or the National Institute of Education.

Description of Evaluation Report Series

The Comprehensive School Mathematics Program (CSMP) is a program of CEMREL, Inc., one of the national educational laboratories, and is funded by the National Institute of Education. Its major purpose is the development of curriculum materials for grades K-6.

Beginning in September, 1973, CSMP began an extended pilot trial of its Elementary Program. The pilot trial is longitudinal in nature; students who began using CSMP materials in kindergarten or first grade in 1973-74, were able to use them in first and second grades respectively in 1974-75, and so on in subsequent years. Hence the adjective "extended".

The evaluation of the program in this extended pilot trial is intended to be reasonably comprehensive and to supply information desired by a wide variety of audiences. For that reason the reports in this series are reasonably non-technical and do not attempt to widely explore some of the related research issues. The list of reports from the first two years of the extended pilot trial is given on the next page. The most comprehensive of these are the following:

1-A-1: Overview, Design and Instrumentation
1-A-3: Final Summary Report, Year 1
2-A-1: Final Summary Report, Year 2
3-B-1: Summary of Second and Third Grade Test Data Year 3
4-A-1: Final Summary Report, Year 4
5-B-1: Fourth Grade MANS Test Data
and 6-A-1: Summary of Evaluation Data Through Fourth Grade

Longitudinal Pilot Study of the
Comprehensive School Mathematics Program

Evaluation Report Series

Evaluation Report 1-A-1	Overview, Design and Instrumentation
Evaluation Report 1-A-2	External Review of CSMP Materials
Evaluation Report 1-A-3	Final Summary Report Year 1
Evaluation Report 1-B-1	Mid-Year Test Data: CSMP First Grade Content
Evaluation Report 1-B-2	End-of-Year Test Data: CSMP First Grade Content
Evaluation Report 1-B-3	End-of-Year Test Data: Standard First Grade Content
Evaluation Report 1-B-4	End-of-Year Test Data: CSMP Kindergarten Content
Evaluation Report 1-B-5	Test Data on Some General Cognitive Skills Related to CSMP Content
Evaluation Report 1-B-6	Summary Test Data: Detroit Schools
Evaluation Report 1-C-1	Teacher Training Report
Evaluation Report 1-C-2	Observations of CSMP First Grade Classes
Evaluation Report 1-C-3	Mid-Year Data from Teacher Questionnaires
Evaluation Report 1-C-4	End-of-Year Data from Teacher Questionnaires
Evaluation Report 1-C-5	Interviews with CSMP Kindergarten Teachers
Evaluation Report 1-C-6	Analysis of Teacher Logs
Evaluation Report 2-A-1	Final Summary Report Year 2
Evaluation Report 2-B-1	Second Grade Test Data
Evaluation Report 2-B-2	Readministration of First Grade Test Items
Evaluation Report 2-B-3	Student Interviews
Evaluation Report 2-C-1	Teacher Questionnaire Data
Evaluation Report 2-C-2	Teacher Interviews, Second Grade
Evaluation Report 2-C-3	Teacher Interviews, First Grade
Evaluation Report 3-B-1	Second and Third Grade Test Data Year 3
Evaluation Report 3-C-1	Teacher Questionnaire Data Year 3
Evaluation Report 4-A-1	Final Summary Report Year 4
Evaluation Report 4-B-1	Standardized Test Data, Third Grade
Evaluation Report 4-B-2	Mathematics Applied to Novel Situations (MANS) Test Data
Evaluation Report 4-B-3	Individually Administered Problems, Third Grade
Evaluation Report 4-C-1	Teacher Questionnaire Data, Third Grade
Evaluation Report 5-B-1	Fourth Grade MANS Test Data
Evaluation Report 5-B-2	Individually Administered Problems, Fourth Grade
Evaluation Report 5-C-1	Teacher Questionnaire and Interview Data, Fourth Grade
Evaluation Report 6-A-1	Summary of Evaluation Data Through Fourth Grade
Evaluation Report 6-B-1	Comparative Test Data: Fourth Grade
Evaluation Report 6-B-2	Preliminary Test Data: Fifth Grade
Evaluation Report 6-C-1	Teacher Questionnaire Data: Grades 3-5

Key to Indexing

Evaluation Reports are labelled m-X-n,

where m is the year of the pilot study, with 1973-74 as Year 1.
X is the type of data being reported where A is for overviews and summaries, B is for student outcomes and C is for other data.
n is the number within a given year and type of data.

Preliminary Test Data: Fifth Grade

Table of Contents

Summary 1

Introduction 3

Setting: Selection of Classes 5
The MANS Tests 7

Comparison of Class Mean Scores 9

Discussion of the Results 17

Other Findings: Comparison with Fourth Grade MANS . . 19
Scores by Ability Level 21
Scores by Sex 23
Key Test Items 25
Correlational Data 28
Total MANS Scores 29

Appendix A: Scale-by-Scale Statistical Analysis . . . 31

Appendix B: Reproduction of Student Pages 75

Summary

This is a report of the first year of a two-year evaluation of the fifth grade materials of the Comprehensive School Mathematics Program (CSMP). In preparation for next year's full-scale evaluation of fifth grade, some student achievement scales were administered to a small group of St. Louis area classes, both CSMP and Non-CSMP. The study was intended to be both a pilot test of certain scales which might be revised for next year, and a preliminary comparative evaluation of CSMP/Non-CSMP achievement.

Briefly the main results are as follows:

1. There tended to be significant differences in favor of CSMP classes on scales dealing with mental arithmetic, number relations, number systems (other than whole numbers), certain word problems and problems with multiple correct answers.

2. There tended to be no significant differences on scales dealing with standard computation, geometry, probability and place value.

3. Similar results were obtained when CSMP-Non-CSMP comparisons were made at different levels of reading ability and by sex.

4. The above results are very consistent with previous findings, particularly these obtained in a large scale comparison of fourth grade classes. (See Evaluation Report 6-B-1).

Appendix A provides detailed statistical data about the scales, and items within the scale.

Introduction

In the fall of 1973, the Comprehensive School Mathematics Program (CSMP) began an extended (longitudinal) pilot trial of its Elementary School Program. Over 100 teachers began using the program, either in first grade or kindergarten. In subsequent years some of these classes continued with CSMP materials in the next grade levels, and many new classes began using CSMP materials. The 1978-79 school year provided that portion of the Extended Pilot Trial (EPT) consisting of an extensive evaluation of the CSMP fourth grade materials and a preliminary evaluation of the fifth grade materials (the extensive evaluation of the latter coming in 1979-80.)

This is the report on the preliminary evaluation of the fifth grade materials. It describes the results of administering a series of tests to the 9 fifth grade CSMP classes and to 8 comparison classes in just the St. Louis area. The tests are collectively called the MANS Test (Mathematics Applied to Novel Situations), and are intended to compare CSMP and Non-CSMP students on a variety of mostly novel problem situations, although some of the test scales consist of ordinary computation problems. It is hoped that many of the thinking skills underlying the scales reflect those embodied in many of the CSMP lessons and hence one hopes these scales can be viewed as transfer tasks.

Other Evaluation Reports describe the administration of MANS Tests at lower grade levels; fourth grade (Evaluation Reports 5-B-1 and 6-B-1), third grade (Evaluation Report 4-B-2) and second grade (Evaluation Report 3-B-1).

Setting

Selection of Classes

The 1978-79 school year was the first year of the Extended Pilot Test of fifth grade CSMP materials (Mathematics for the Intermediate Grades, Part III and IV). Altogether there were 33 classes located in 12 districts. The St. Louis area was the only one in which there were sufficient CSMP and comparison classes to make a study of this kind worthwhile. In the St. Louis area nine fifth grade CSMP classes, located in the Normandy, Maplewood-Richmond Heights and Ladue school districts and St. Timothy's parochial school were compared with 8 Non-CSMP classes, selected from adjacent schools where possible, but in any case with the intent that this group of students, teachers and schools be as similar to the CSMP group as possible. The CSMP classes were all "veteran" classes, i.e. had studied CSMP from at least the beginning of fourth grade.

In conjunction with the MANS testing (during two days in the month of May) the Reading Comprehension section of the Stanford Achievement Test was administered to each of the 17 classes. Thus it was possible to compute the mean score for each class and to use this mean as a covariate in the subsequent analysis of class means. It was also then possible to compare CSMP and Non-CSMP students at different levels of reading ability, as measured by that section of the SAT.*

*The SAT Reading Comprehension test was divided into equivalent halves. Random halves of each class took one or the other of the two halves of the test. This was done in the interest of conserving testing time. The score used in all cases was the total number correct on whichever half of the test was taken.

Only those students who were present when the SAT reading section was administered were included in the analysis. This excluded those students who happened to be absent that day. Students joining the class after the month of March were also excluded. As a result of these two types of exclusions, an average of about one student per class was excluded. The average remaining number of students per class was about 22 for CSMP and 20 for Non-CSMP classes.

The mean covariate score across CSMP and Non-CSMP classes turned out to be very close: 23.9 for CSMP classes versus 23.2 for Non-CSMP. The distribution of class means is given in Table 1 below, which lists the approximate percentile ranks of each class.

Table 1
Percentile Ranks, SAT Reading Comprehension Section

CSMP Classes	Non-CSMP Classes
81	82
79	76
76	74
60	65
57	50
57	48
56	47
47	30
44	

Table 1 shows that the corresponding percentile ranks are fairly close, except for the lowest scoring Non-CSMP class. Table 1 also shows that the median percentile rank was about 60 for both CSMP and Non-CSMP classes. Thus this group of classes was somewhat above average in ability.

The MANS Tests

The MANS (Mathematic Applied to Novel Situations) tests consist primarily of specially developed test scales designed for the purpose of comparing CSMP and Non-CSMP students' ability to deal with certain kinds of novel mathematical situations. The scales are intended to be related so some of the thinking processes which seem to occur in the CSMP curriculum, but without containing the language or content of the curriculum.

Previous MANS Tests were developed for use at second, third and fourth grade levels, and the present ones build on those previous efforts. The development of the various scales is a fairly lengthy process and includes extensive reviews (one group of reviewers being the external CSMP Evaluation Panel) and pilot tests. Appendix A of Evaluation Report 4-B-3 describes in detail how this process worked with the third grade MANS tests; the process was similar for the present scales.

Since the 1978-79 school year was the first of two years for the EPT at fifth grade, the MANS test used in testing fifth graders was something of a trial version. Some of the scales were old scales developed carefully for the fourth grade level and some were new in the process of being pilot tested. In addition, some computation scales were included which were not in any way novel, in order to compare classes on straightforward computation skills (since the various districts involved did not use a common standardized test from which this type of comparative data could be obtained).

Thus the full range of MANS scales normally given in the second EPT year was not given here, in order that some new scales could be developed. This fact, together with the small number of classes involved means that the results, especially concerning total MANS scores, must be viewed rather tentatively at this time.

Altogether, 19 scales were used, excluding the one for the covariate reading score. For each scale a specially trained tester explained what the required task was and did one or two samples with the whole class, all according to a carefully designed script. Then the students went ahead and did the test items on that scale, which involved one or more pages. They did not turn the page to begin a new scale until they were told to do so by the tester - then everyone turned at once - and the same procedure was followed in the next scale. The typical scale consisted of 7-11 items and required about 6 or 7 minutes (though there was usually no fixed time limit).

For certain scales with identical directions, alternate forms were given to random groups of studnets. This item sampling procedure increased the number of items and scales on which class mean scores could be estimated for the limited testing time that was available.

Comparison of Class Mean Scores

In order to simplify the reporting somewhat, similar test scales will be grouped together, described briefly and the results given. This grouping is somewhat arbitrary, but should make it easier for the reader to keep track of the various scales.

Appendix A gives data for the complete set of scales, one after the other, with two pages devoted to each scale. The test items are listed, together with a great deal of statistical data including item analyses, graphs of class means and Analysis of Covariance data. In fact Appendix A by itself constitutes a rather full report of the MANS testing.

On the next few pages, each scale is described briefly and the comparison of class mean scores is presented. The unit of analysis is the class. For each portion of the MANS Tests, a mean score is computed for each class. Then the two sets of class means (for CSMP and Non-CSMP) are compared statistically, using an analysis of covariance procedure to determine if the between-curricular groups differences are greater than might have occurred by chance, given the existing differences in reading scores.

Due to the small number of classes involved a fairly large difference between the two curricular groups is needed to register as statistically significant. Since this is merely a preliminary evaluation of the fifth grade materials, these inherent difficulties should not be too troublesome.

SCALE CATEGORY: COMPUTATION

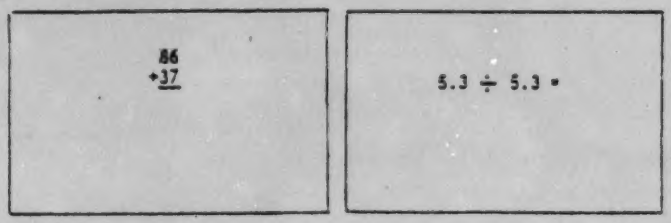
	Adjusted Means[1] CSMP	Non-CSMP	p-value less than:[2]
(C7) Standard Computation (13 items in each of three forms)	25.3	25.1	.99

(Students took one of three randomly-distributed alternative forms.)

Abstract:

Typical computation items were selected from various standard achievement tests for fifth grade. Of the 39 items involved in the three forms 23 involved only whole numbers, 8 fractions, and 8 decimals; while 10 involved addition, 8 subtraction, 12 multiplication, and 9 division.

Samples:



(C3) Mental Arithmetic: Addition	3.9	3.6	.10

Abstract: An open number sentence involving addition must be completed without aid of pencil and paper. 5 items

Sample: 53 + 8 = ☐

(C4) Mental Arithmetic: Subtraction	3.6	3.0	.04

Abstract: Same as C3, but with subtraction, 5 items.

Sample: 72 − 4 = ☐

(C5) Mental Arithmetic: Multiplication	5.9	5.5	.15

Abstract: Same as C3, but with multiplication, 8 items

Sample: ☐ x 4 = 400

(C6) Mental Arithmetic: Division	5.4	4.4	.03

Abstract: Same as C3, but with division, 8 items.

Sample: 150 DIVIDED BY 25 = ☐

Total Computation	44.1	41.2	.29

[1] These are class means adjusted for differences in ability as measured by the reading comprehension score from the Stanford Achievement Test; CSMP classes had slightly higher reading scores so their MANS scores would be adjusted downward while the Non-CSMP scores would be adjusted upward. However, these adjustments were seldom greater than 0.1.

[2] This is the p-value of the resulting F test from the Analysis of Covariance on the classs means with 1 and 14 degrees of freedom. Low p-values are sometimes said to be "significant" if they are less than say .05; i.e., there is a 5% possibility that the observed differences between CSMP and Non-CSMP classes could have happened merely by chance had there been no differential curriculum effects.

	Adjusted Means[1] CSMP	Non-CSMP	p-value less than:[2]

SCALE CATEGORY: OTHER NUMBER SYSTEMS

(N1) Decimal Gas (7 items) 3.9 2.8 .02

(Taken by the random 1/2 of the students not taking W3)

Abstract: With word problems about gasoline, one step solutions are required in which the numbers involved are in decimal form.

Sample:

Tom has 6.5 gallons.
He buys 3.5 more gallons.
How much gas will he have then? _____

(N2) Negative Hits and Misses (6 items) 3.4 2.9 .07

(Taken by all students)

Abstract: Given two rules ((a) each hit means a gain of 5 points (b) each miss means a loss of 1 point) and given a vertical number line running from 12 below zero to 15 above, players turns are described in part with the required task being to complete the description.

Sample:

Peter	Started with a score of	Number of Hits	Number of Misses	Ended with a score of
	10 below zero	1		12 below zero

Total, Other Number Systems 7.2 5.9 .03

	Adjusted Means[1]		p-value less
	CSMP	Non-CSMP	than:
SCALE CATEGORY: ESTIMATION			
(E5) Word Problem Approximations (6 items)			
(Taken by all students)	3.1	2.8	.05

Abstract: Given 15 seconds (for each item) one of four round-number answers must be chosen which is closest to the exact answer to a word problem (read aloud by the tester) with relatively easy calculations.

Sample:

> Susan has $131.
> Chairs cost $32.
> About how many chairs can Susan buy?
>
> 2 chairs 4 chairs 6 chairs 10 chairs

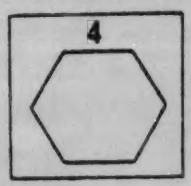
SCALE CATEGORY: GEOMETRY

(G1) Geometric Congruencies (8 items)

(Taken by all students) 5.8 5.5 .36

Abstract: The students first examine 3 correct and 3 incorrect solutions to a sample problem. Then, given a regular geometric shape and a number, the shape must be divided into that many congruent parts.

Sample:

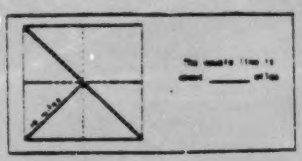


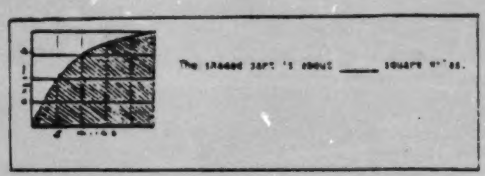
(G2) Geometric Approximations (7 items)

(Taken by all students) 2.3 2.0 .19

Abstract: Consists of two sections (pages). On the first page, three geometric figures are presented with the distances of some lines given, Student must derive an approximate distance for another designated line. On the second page, four geometric figures are presented with partial areas or distances given. Student must derive an approximate area of another part of the figure.

Samples:





Total Geometry 8.0 7.6 .14

SCALE CATEGORY: PROBABILITY

	Adjusted Means[1] CSMP	Non-CSMP	p-value less than:[2]

(P1) Spinners (19 "items")

(Taken by all students) 7.4 7.5 .94

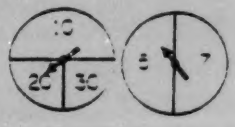
Abstract: On page one, one spinner is shown and the student is told to imagine it being spun 100 times. Than a series of five questions are asked: "about how many times would we get a score of _____?"

Sample:

SUPPOSE WE PLAY THE GAME 100 TIMES

About how many times would we get a score of less than 4? _____

Abstract: On page two, two spinners are shown with a similar circumstance imagined and a similar 5 questions asked except that the score of the two spinners are to be added. Maximum score is 19 due to credit for relatively correct answers.



Sample:

SUPPOSE WE PLAY THE GAME 100 TIMES

About how many times would we get a score of 25? _____

(P2) Which Is More Likely (11 items)

(Taken by all students) 5.5 5.6 .84

Abstract: Three different situations are posed and for each the student is asked to compare the likelihood of 3 (or 4) pairs of outcomes. Student is to circle the more likely of the pair or circle both, if equally likely.

Sample:

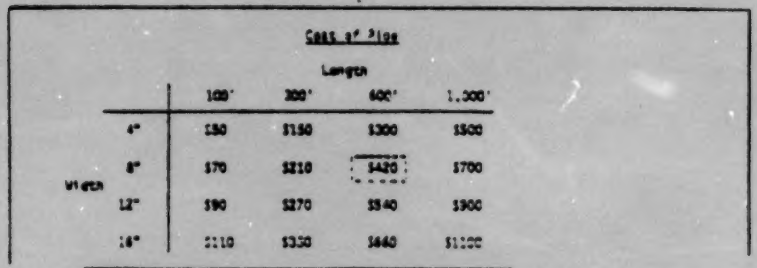
	Which is More Likely		
Roll a pair of dice	this? ↓	or	this? ↓
	total of 7		total of 11

Total Probability 12.9 13.1 .85

	Adjusted Means[1] CSMP	Non-CSMP	p-value less than:[2]
SCALE CATEGORY: NUMBER RELATIONS			
(R6) Use of Data (7 items)	2.9	2.6	.06

(Taken by all students)

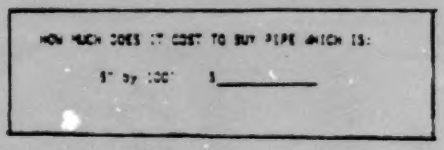
Abstract: Students are instructed as to how the "Cost of Pipe" table is read (see page reproduced below). The first two sample items are done as a class with much tester help followed by the students doing the practice item. These first three items are all possible by just reading directly from the table. Then students are shown that all remaining items can't be taken directly from the table but must be estimated (by interpolation or extrapolation). The sample item is explained by the tester. Students then do the last 7 items on their own.



Cost of Pipe

Width \ Length	100'	200'	600'	1,000'
4"	$50	$150	$300	$500
8"	$70	$210	$420	$700
12"	$90	$270	$540	$900
18"	$110	$330	$660	$1100

Sample:



HOW MUCH DOES IT COST TO BUY PIPE WHICH IS:

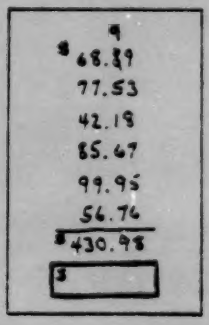
8" by 100' $_____

(R7) Using Hints (10 items) 3.3 2.6 .08

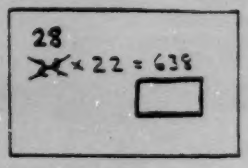
(Taken by all students)

Abstract: Students are told that an excellent math student has done the items (all correctly) except that he always copies the items wrong. The copying correction is pointed out and the first item worked with the testers instructions. Tester stresses that there is limited time (exactly 4 minutes) and one should not recalculate the items, simply change the previously correct answer to compensate for the copying error. 7 addition and 3 multiplication items.

Samples:



$68.39
77.53
42.19
85.47
99.95
56.74
$430.98
$



28
× 22 = 638

Total Number Relations 6.2 5.2 .05

	Adjusted Means[1] CSMP	Non-CSMP	p-value less than:[2]

SCALE CATEGORY: ELUCIDATION

(U1) Elucidation (4 items, maximum of 24 points) 12.5 9.9 .02

(Taken by all students)

Abstract: For each item a situation is posed and a question asked orally by the tester in some detail and briefly on the student page. The suggestion is that there are several correct answers to the question. One of these is given. Student is to give all the others. Student's score is total number of correct answers minus total number of incorrect answers.

Sample:

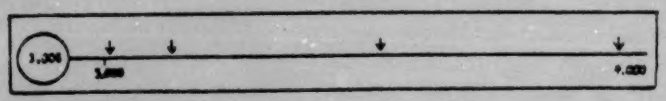
Start at zero
Get to 24 on the nose.
Counting by 2.

SCALE CATEGORY: PLACE VALUE

(V1) Arrowed Number Lines (11 items) 5.9 5.8 .85

Abstract: Given a segment of a number line with four little arrows pointing to four different points, circle the arrow pointing closest to a given number. Last three items involve fractions.

Sample:



SCALE CATEGORY: WORD PROBLEMS (See also E5, N1, N2, L1 each of which use a word problem format but have more salient characteristics and hence are catagorized differently.)

(W3) Three Stage (5 items) 2.0 1.5 .07

(See N1)

Abstract: Student must read a 3 to 5 sentence word problem and complete a solution involving at least three operations.

Sample:

Mary has 4 more marbles than Pete.
Pete has 2 more marbles than Lisa.
Lisa has 3 more marbles than Ed.
If Mary has 20 marbles, how many does Ed have? _____

Discussion of the Results

In order to facilitate a discussion of the foregoing results, Table 2 gives a summary by scale category.

Table 2
Summary of Results by Scale Category

	Adjusted Means		P-value
	CSMP	Non-CSMP	less than .05
Computation (C7,C3,C4,C5,C6)	44.1	41.7	.29
Estimation (E5)	3.1	2.8	.05
Geometry (G1,G2)	8.0	7.6	.14
Other Number Systems (N1,N2)	7.2	5.9	.03
Probability (P1,P2)	12.9	13.1	.85
Number Relations (R6,R7)	6.2	5.2	.05
Elucidation (U1)	12.5	9.9	.02
Place Value (V1)	5.9	5.8	.85
Word Problems (W3)	2.0	1.5	.07
Total of all scales	103.5	91.7	.07

At least three important features should be pointed out in connection with Table 2.

1. On most categories of scales the results are in the main quite consistent with previous findings and with relative emphases of CSMP and traditional curricula. CSMP students have usually scored well on scales with the categories of estimation, other number systems, number relations, word problems and elucidation problems with multiple correct answers). Similarly it is not surprising to see no CSMP-non-CSMP difference on the place value scale.

2. Combining the various computation scales into an omnibus category obscures an important finding. There were really two distinct kinds of computation items: straightforward computation (on which CSMP-non-CSMP scores were virtually identical) and mental arithmetic (on which CSMP classes had a significant overall advantage.) This finding is also in agreement with what one would expect based on previous data and on the nature of the CSMP curriculum.

3. Somewhat more puzzling is the lack of significant CSMP-Non-CSMP differences in the areas of geometry and probability. Given the relatively greater emphasis on these two topics in the CSMP curriculum, one would expect to see CSMP students doing better. However, there is some evidence from questionnaire data that indicates that the CSMP geometry and probability lessons are not as faithfully taught as are those in other areas. Attention will no doubt be paid to a clsoer look at these two areas in the extensive evaluation of fifth grade during 1979-80.

Other Findings

Comparison with Fourth Grade MANS

Some of these scales were also administered to fourth grade classes as part of a larger comparison study (see Evaluation Report 6-B-1). The table below gives a rather general comparison of 4th and 5th grade results by scale category.

Table 3

Comparison of Fourth & Fifth Grade Results (1979)

Common Scale Categories	No. of Scales		No. of Scales in common	Probability Level of the Difference					
				4th Grade Differences in Favor of:			5th Grade Differences in Favor of:		
	4th	5th	4th & 5th	CSMP	Non-CSMP	NSD	CSMP	Non-CSMP	NSD
Computation	8	7	4			*			*
Estimation	5	1	1	.01			.05		
Geometry	1	2	1			*			*
Other Number System	3	2	2	.01			.03		
Probability	1	2	1	.01					*
Number Relations	5	2	0	.01			.05		
Place Value	2	1	1			*			*
Word Problems	2	1	0	.01			.07		

*NSD = no significant difference (p-value $\geq$ .13)

Apart from the uniformity of results in favor of CSMP, perhaps the most noteworthy feature of Table 3 is the similarity of results between fourth and fifth grades. For four of the categories (Estimation, Other Number Systems, Number Relations and Word Problems), there were sizable differences in favor of CSMP for both years. In each case the difference was statistically more significant in fourth grade than in fifth, probably due in large part to the larger number of scales in each category in fourth grade. Similarly in three other scale categories (Computation, Geometry and Place Value), there were non-signficant differences in both grades. The one exception is the case of the Probability category: a significant difference at the fourth grade in favor of CSMP and essentially no difference at fifth grade.

Scores by Ability Level

Based on covariate scores, all students were put in rank order and divided into approximately equal groups. These quarter groups (Q1,Q2,Q3 and Q4) were again separated into subgroups of CSMP and Non-CSMP. This resulted in about 48 CSMP students and 42 Non-CSMP students in each group.

Within each ability group, mean scores were derived on each MANS test. Figure 1, below, shows a graph of the total of these means across all the MANS scores plotted against the mean reading score for each ability group.

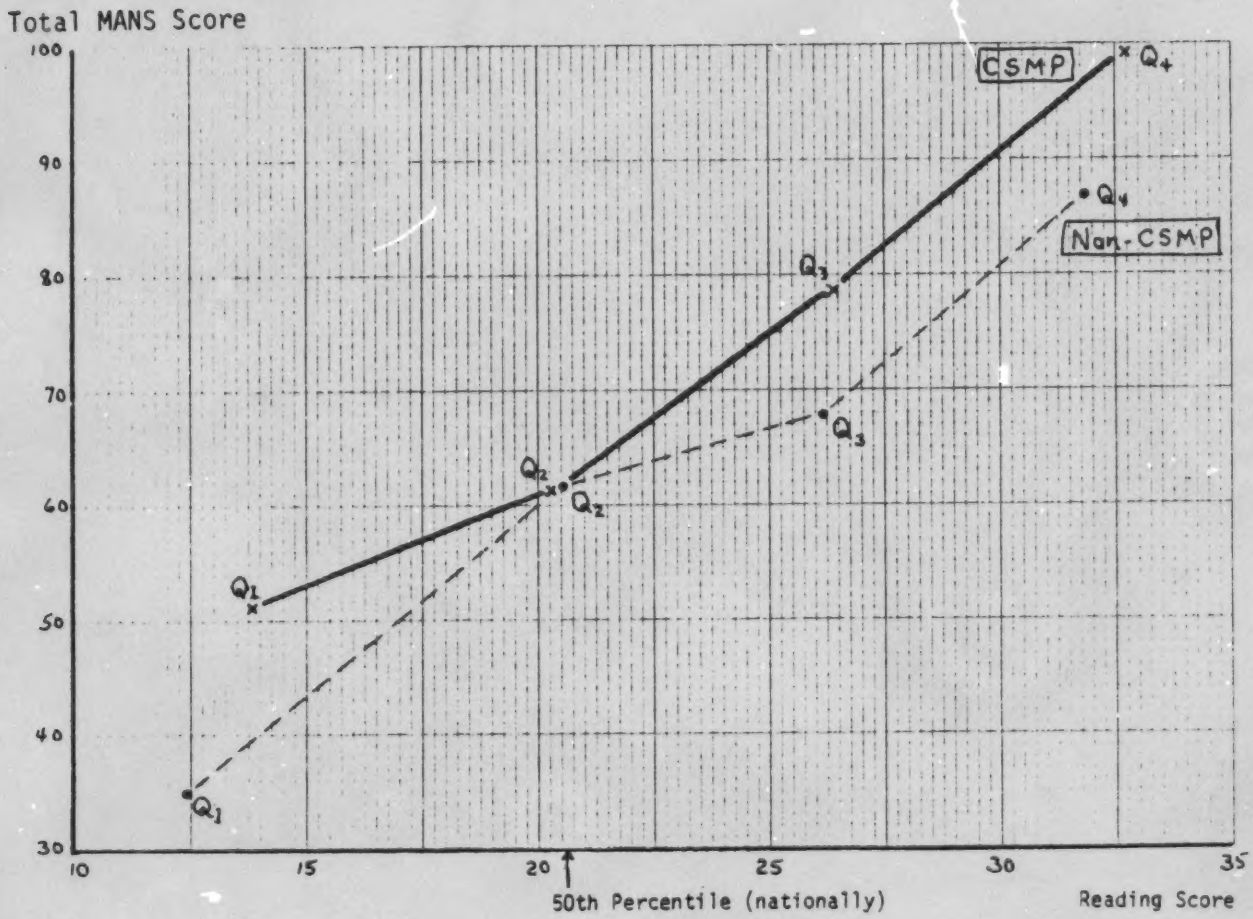
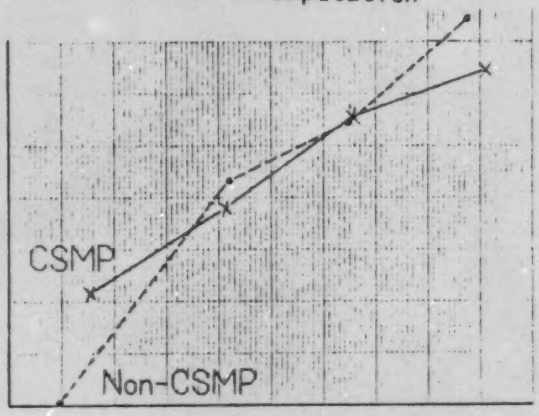


Figure 1 Means by Ability Groups on Total MANS Score

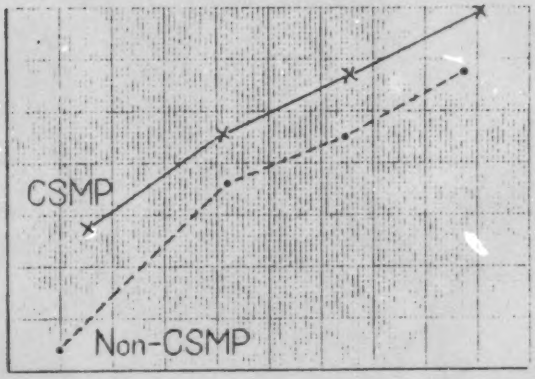
Figure 1 shows that the CSMP superiority on the MANS test was fairly consistent except at the second quartile where the mean scores were equal. (Note also that the 50th percentile (nationally) on the reading score corresponds to the second quartile means. Again caution is advised in making any generalizations from this data because of the small numbers involved. For example, about half the Non-CSMP students in the lowest quartile came from the same class.

Figure 2 shows the abbreviated graphs for the two types of computation scales, illustrating again the conclusion that there is very little difference on straightforward computation while CSMP students do considerably better in mental arithmetic.

C7: Standard Computation



C3-6: Mental Arithmetic



Means by Ability Group on Two MANS Subscores

This same procedure was used with each individual scale and the results are shown in the scale-by-scale analysis in Appendix A. Generally the results were consistent across ability levels.

Scores by Sex

Separate means were calculated for boys and for girls. This was done separately for CSMP and Non-CSMP students and then combined. This information is summarized in Table 4.

Table 4
Mean MANS Scores by Sex

Scale Category	Number of Scales	Sex	CSMP	Non-CSMP	Combined	Significance=.01 (Boys combined versus girls combined)
Computation	7	Boys	44.7	40.5	42.8	
		Girls	44.3	40.7	42.6	
Estimation	1	Boys	3.3	2.8	3.1	
		Girls	3.0	2.6	2.8	
Geometry	2	Boys	8.2	7.4	7.9	
		Girls	7.9	7.2	7.6	
Other Number Systems	2	Boys	8.1	6.1	7.2	*
		Girls	6.3	4.5	5.4	
Probabilty	2	Boys	14.2	13.3	13.8	*
		Girls	11.6	11.4	11.5	
Number Relations	2	Boys	6.5	5.1	5.9	
		Girls	5.8	4.4	5.1	
Elucidation	1	Boys	13.3	9.6	11.6	
		Girls	12.0	9.1	10.6	
Place Value	1	Boys	6.4	5.9	6.2	*
		Girls	5.5	5.1	5.3	
Word Problems	1	Boys	2.2	1.6	1.9	
		Girls	2.1	1.3	1.7	
Grand Total Mans	14	Boys	76.7	63.9	71.0	*
		Girls	67.2	59.4	63.6	
SAT Reading	-	Boys	24.2	22.1	23.3	
		Girls	23.1	23.5	23.3	
Number		Boys	108	88	196	
		Girls	85	76	161	

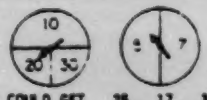
Note that the grand total is based on 14 scales, the number any one student took. Altogether there were 19 scales invovled in alternate forms of the test. Hence the figures for Grand Total will be less than the sum of all the separate scales.

It can be seen that when comparing CSMP versus Non-CSMP, the results were consistent for each sex. CSMP boys and girls both did better than their Non-CSMP counterparts, and usually by about the same amount, if one takes into account the higher reading ability of the CSMP boys. When boys were compared with girls overall (the right hand column of Table 4), it can be seen that, while the reading and computation scores were the same, boys had an advantage on all other categories, significantly so in several cases. Further analysis of only the Total MANS scores, shows that boys did significantly better than girls at the lowest two reading levels, but were about the same at the highest two levels. Unfortunately, at this level of analysis, the number of students in each cell (about 45) has become rather small and one must be cautious about further generalizations.

Key Test Items

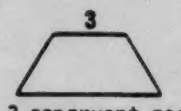
It is interesting to note which test items produced the largest differences in percent correct between CSMP and Non-CSMP students. One can take into account the slightly higher ability scores attained by CSMP students on each item by making a generous rule of thumb adjustment of 7% in favor of Non-CSMP students on each item.

There were 10 items on which Non-CSMP students did clearly better. These items are shown below; on all these items Non-CSMP students did at least 13 adjusted percentage points better than CSMP students.

Scale	Item
Contained in one of the alternate forms of C1	$\begin{array}{r} 12604 \\ -8958 \\ \hline \end{array}$
	$5\overline{)255}$
	$13\overline{)273}$
	$\frac{3}{8} + \frac{1}{4}$
	$\frac{1}{2} + \frac{1}{3}$
	$\frac{2}{6} + \frac{3}{6}$
	$\frac{4}{5} + \frac{3}{5}$
E5: Word Problem Approximations	Tom and Bill had a long distance race. Bill ran it in 32 minutes, 2 seconds. Tom ran it in 29 minutes, 34 seconds. About how much did Tom win by? $1\frac{1}{2}$ minutes $2\frac{1}{2}$ minutes $3\frac{1}{2}$ minutes $4\frac{1}{2}$ minutes
P1: Spinners	 SCORES YOU COULD GET 25, 17, 35, 15, 27, 37 (Spin both spinners and add the two results) About how many times would we get a score of 17? ___ About how many times would we get a score of 25? ___

It is worthwhile noting that four of the ten items dealt with the addition of fractions and two dealt with the division algorithm. No doubt these results reflect the lack of emphasis in the fifth grade CSMP materials on these computational skills. Not so easily explained is the Non-CSMP superiority on the two probability (spinners) items, especially since most elementary programs put little or no emphasis on probability while the CSMP materials include lessons in this area form the earliest grades.

There were 16 items on which CSMP students did clearly better than Non-CSMP students (better by at least 13 adjusted percentage points):

Scale	Item
	100 -79
	$\frac{1}{3}$ of 36
Contained in one of the alternate forms of C7	0.21 x5
	1.5 +3.7
C6 Mental Arith. (÷)	500 ÷ 2 = ☐ 150 ÷ 25 = ☐ 800 ÷ ☐ 200 ☐ ÷ 3 = 120
C5 Mental Arith (×)	200 x ☐ = 800
E5 Word Number Approximations	George has $902. He wants to buy a piano costing $1,348. About how much more money does he need? $250 $350 $450 $550
G1 Geometric Congruencies	 3 in 3 congruent parts

Scale	Item
N1 Decimal Gas	Bill has 6.5 gallons. He buys another half gallon. How much gas will he have then? Tom has 6.5 gallons. He buys 3.5 more gallons. How much gas will he have then? Ron has 6.5 gallons. Next week he will use ten times this much. How much gas will he use next week? Ken has 6.5 gallons of gas. He gives away half of it. How much gas will he have left?
R7 Using Hints	2 44.18 96.17 88.99 75.75 63.32 49.56 418.67

Ten of these items involved either the a) process of division or b) numbers in decimal form. Since CSMP emphasizes decimals early on, result b) is not too surprising. Also since the items concerning result a) involved relatively small numbers, an understanding of the process of division was more salient than an ability to apply the standard algorithm.

Correlational Data

Three types of correlational data will be noted briefly. First, the KR 20 Reliability/Homogeneity coefficients were calculated for each scale for each curricular group and are reported in Appendix A. This index is a measure of the internal consistency of a scale, i.e., the degree to which the items are testing basically one underlying ability. For all but 3 of the scales this coefficient was above .50, and the median value was .70 which is fairly high considering the relative brevity of most scales, (an average of only about 8½ items each.) Each of the three exceptions were very short scales; also one was very easy and another very difficult for the students.

Second, correlations between each scale and the SAT Reading Test were calculated and these too are reported in Appendix A. With the exception of one scale these correlations fell within a range of .31 to .69 with an average of .48. In comparison with other years. these correlations are lower. Perhaps this can be explained in large part by the fact that the covariate this year was strictly a reading test compared to a more general test of intelligence used in the past. It is noteworthy that the average of these correlations for the CSMP group of students (.43) was much lower than the average for Non-CSMP (.53). Perhaps this lends further support for the theory that CSMP taps a non-traditional mode of thinking.

The third type of correlations had to do with intercorrelations between MANS TESTS. Of the 163 meaningful intercorrelations, the median was about .5, all but 11 falling in the range .30 - .70. The two scales that correlated best with the others were C6 (Mental Arithmetic: Division) and N1 (Decimal Gas). The two scales that correlated worst were P2 (Which is More Likely) and one of the C7's (Standard Computation).

Total MANS Scores

For each class a total MANS score was calculated by adding together all the mean scores of the individual scales. These means were then plotted against the mean SAT Reading scores and the graph is shown on the next page together with the corresponding regression line. (Appendix A presents similar data for each individual scale.)

From the graph, one can see a difference in favor of CSMP but it is not clear-cut, reflected by the probability of chance occurrance being 7%. As always there is a problem in interpreting what the total MANS score represents. Obviously it is some kind of measure of mathematical abilities and skills, but whenever a large number of somewhat dissimilar scales like these is combined to form a total score, the labelling of this total score is a problem.

This is particularly true in this case. Since this was partially a pilot test of some new scales, it was not possible to also have the range of tests one would like (but which <u>will</u> be possible in next year's larger scale comparisons). Hence, the Total MANS score does not include several scales which have historically produced large differences in favor of CSMP classes - see the fourth grade scales in Evaluation Report 6-B-1.

Thus, it is probably more useful to attend more carefully to the results from the various individual scales, than those from any total score.

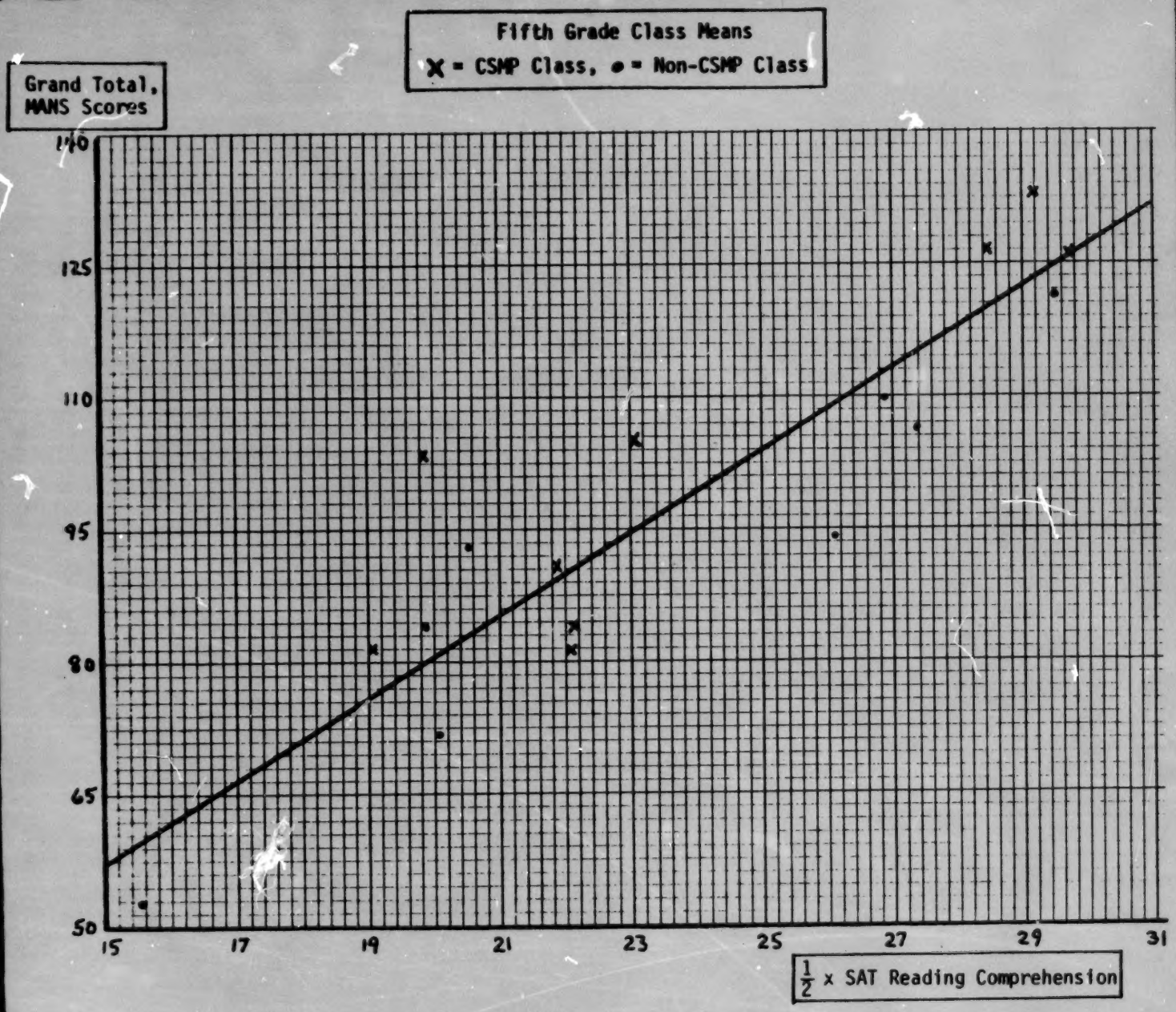


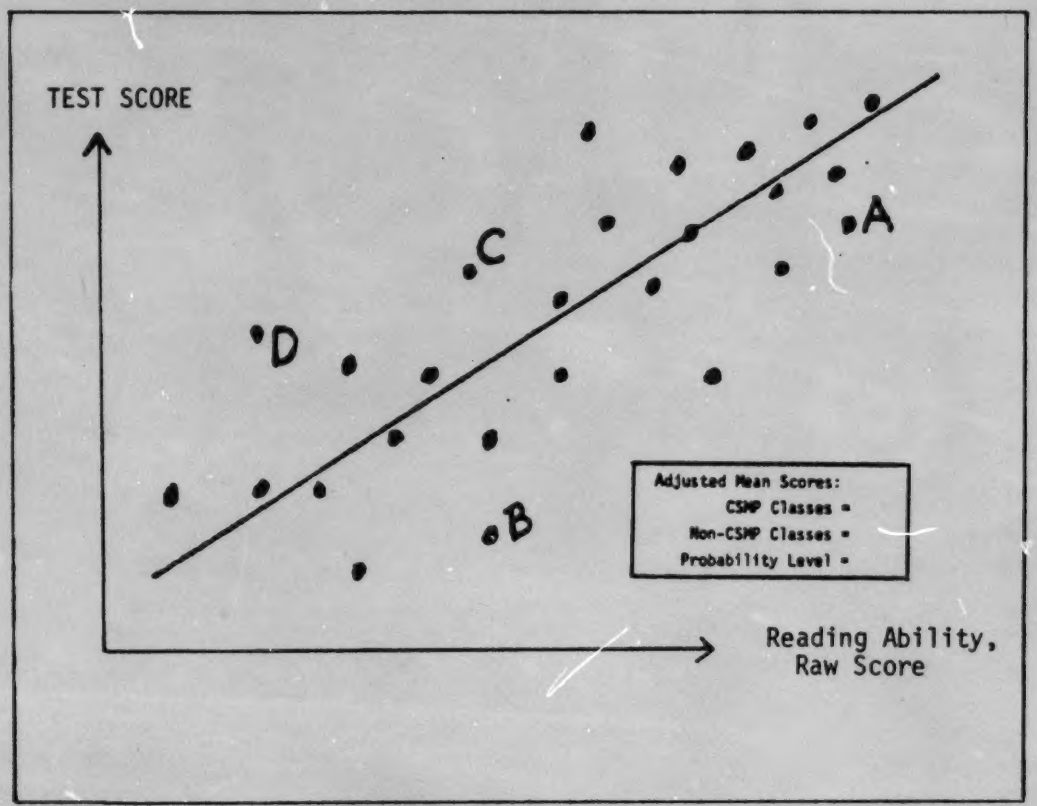
Figure 3: Class Means for MANS Total Score vs. Reading Score

Appendix A

Scale-by-Scale Statistical Analysis

This Appendix contains information on each MANS scale. There are two pages for each scale. The left hand page provides numerical data, especially item analyses, based on a summary across students. The right hand page is a graph showing mean scores of each class.

Right Hand Pages



The right hand pages are for the graphs of class means. Each CSMP class is represented by an x and each Non-CSMP by a dot (•). Based on this set of class means thus graphed, the regression line has been drawn. This line is the best linear prediction of mean class test score that can be made from knowing the ability level of the class. Note in the example that classes A and B fall well below the regression line, or are scoring well below what would be predicted for them from knowing the ability level of the class, while classes C and D fall well above the regression line. Note also that, although class A had a slightly higher mean score on the test than did class C, class C did much better given the relative ability scores of the two classes. When the class means generally fall close to the regression line, test scores are well predicted by the covariate; when they are more dispersed from this line, the covariate is a less effective predictor.

In the box in the lower right hand corner, the mean scores across CSMP and across Non-CSMP classes, adjusted for reading ability, are also given, together with the p-value obtained from the F-test. (The p-value is the probability that a difference in mean scores that large could have occurred by chance alone. Hence, the smaller the p-value (especially below, say, .05), the more likely it is that there are "real" differences between CSMP and Non-CSMP classes.

Test Item	Percent Correct	
	CSMP	Non-CSMP
①		②
Number of Students		③
KR 20 Reliability/Homogeneity		④
Mean Reading Score		⑤
Correlation of Scale & Reading Score		⑥

⑦ Means by Ability Level

	1	2	3	4	Combined
CSMP ($\bar{X}_1$)					
Non-CSMP ($\bar{X}_2$)					
$(\bar{X}_1-\bar{X}_2)/sd$					

⑧ Notes

① The test items listed here are shown in a greatly compressed and sometimes altered form so that what required a full page on the student copy of the test can be squeezed into this space. In those cases in which this reduction did serious damage to the original, a reference is made to Appendix B, where the full student pages of those scales are given.

② These are the percentages of students getting the item correct.

③ The number of CSMP students taking a scale will be either close to 200, close to 100 or close to 65 depending on whether a whole class, a random half of the class, or a random third of the class took the scale. The numbers for Non-CSMP students are a little lower, due to one fewer class.

④ The KR 20 reliability coefficient is a measure of the degree to which the items in a scale are testing a single underlying ability. A large KR 20 (above say, .8) means a high correlation among the items; a low KR 20 (below say .5) means a low correlation among the items and <u>not</u> a single underlying ability.

⑤ The mean raw score on the SAT Reading Test for those students who took this particular scale.

⑥ Further correlation data is given on page 28

⑦ Students were allocated to quartiles based on their Reading scores in such a way that roughly equal numbers of students were in each quartile. This table gives the mean scores for each quartile and for all students combined. The bottom line is the CSMP-Non-CSMP difference, divided by the pooled standard deviation of these means, which has been adjusted to take into account the higher CSMP reading scores - notably at the lowest and highest quartiles. This is a t-statistic, so that a rough rule of thumb would be to consider values of this ratio above 2 to be significant. This is aggregation by <u>students</u> rather than <u>classes</u>, so this is a much more liberal test of differences.

⑧ The notes point out or amplify some of the more important or interesting data.

C3 - Mental Arithmetic +

Test Items	Percent Correct CSMP	Non-CSMP
53 + 8 = ☐	97	88
3,009 + 3,001 + 3,001 = ☐	72	67
542 + 99 = ☐	72	69
☐ + 125 = 250	87	69
989 + ☐ = 1,000	64	62
Number of Students	97	84
KR20 Reliability/Homogeneity	.59	.57
Mean Reading Score	19.7	18.5
Correlation of Scale and Reading Scores	.34	.57

Means by Ability Level

	1	2	3	4	Combined
CSMP ($\bar{X}_1$)	3.2	3.9	4.0	4.5	3.9
Non-CSMP ($\bar{X}_2$)	2.4	3.5	4.0	4.4	3.5
$(\bar{X}_1-\bar{X}_2)/sd$	1.6	1.2	0.1	0.0	1.8

Notes: See page 42 and 43 for notes and graph dealing with all four mental arithmetic scales.

C4 - Mental Arithmetic

Test Items	Percent Correct CSMP	Non-CSMP
72 − 4 = ☐	87	81
1,000 − 5 = ☐	76	62
700 − 401 = ☐	69	51
525 − ☐ = 225	79	64
☐ − 250 = 150	53	35
Number of Students	100	81
KR20 Reliability/Homogeneity	.66	.65
Mean Reading Score	20.0	19.7
Correlation of Scale and Reading Scores	.31	.51

Means by Ability Level

	1	2	3	4	Combined
CSMP ($\bar{X}_1$)	3.0	3.4	3.8	4.2	3.7
Non-CSMP ($\bar{X}_2$)	1.8	2.8	3.1	3.8	2.9
$(\bar{X}_1-\bar{X}_2)/sd$	2.5	1.1	1.8	1.1	3.2

Notes: See page 42 and 43 for notes and graph dealing with all four mental arithmetic scales.

C5 - Mental Arithmetic ✖

Test Items	Percent Correct CSMP	Non-CSMP
7 x 3 = ☐	96	98
7 x 30 = ☐	89	89
3 x 125 = ☐	86	73
12 x 500 = ☐	52	37
99 x 8 = ☐	55	53
☐ x 4 = 400	81	81
200 x ☐ = 800	66	47
☐ x 250 = 500	73	65
Number of Students	100	81
KR20 Reliability/Homogeneity	.71	.69
Mean Reading Score	20.0	19.7
Correlation of Scale and Reading Scores	.35	.42

Means by Ability Level

	1	2	3	4	Combined
CSMP ($\bar{X}_1$)	5.1	5.6	6.2	7.0	6.1
Non-CSMP ($\bar{X}_2$)	4.1	5.7	5.4	6.5	5.4
$(\bar{X}_1-\bar{X}_2)/sd$	1.4	-0.2	1.6	1.5	2.0

Notes: See page 42 and 43 for notes and graph dealing with all four mental arithmetic scales.

C6 - Mental Arithmetic ÷

Test Items	Percent Correct CSMP	Non-CSMP
.21 DIVIDED BY 3 = ☐	94	81
500 DIVIDED BY 2 = ☐	75	55
700 DIVIDED BY 10 = ☐	78	62
300 DIVIDED BY 50 = ☐	62	46
150 DIVIDED BY 25 = ☐	61	42
800 DIVIDED BY ☐ = 200	65	38
☐ DIVIDED BY 3 = 120	51	32
600 DIVIDED BY ☐ = 2	68	56
Number of Students	97	84
KR20 Reliability/Homogeneity	.86	.84
Mean Reading Score	19.7	18.5
Correlation of Scale and Reading Scores	.56	.64

Means by Ability Level

	1	2	3	4	Combined
CSMP ($\bar{X}_1$)	3.2	5.2	6.4	7.1	5.5
Non-CSMP ($\bar{X}_2$)	1.5	4.2	5.5	5.8	4.1
$(\bar{X}_1-\bar{X}_2)/sd$	2.2	1.7	1.3	2.2	4.1

Notes: See page 42 and 43 for notes and graph dealing with all four mental arithmetic scales.

Notes for Scales C3-C6 (Mental Arithmetic):

1. It is interesting to note that CSMP students did relatively best on C4 (Subtraction) and C6 (Division), particularly since these were their weakest operations on the Standard Computation Scale, C7. Again this result may be a reflection of CSMP's greater emphasis on conceptualizing the operations before learning a standard algorithm.

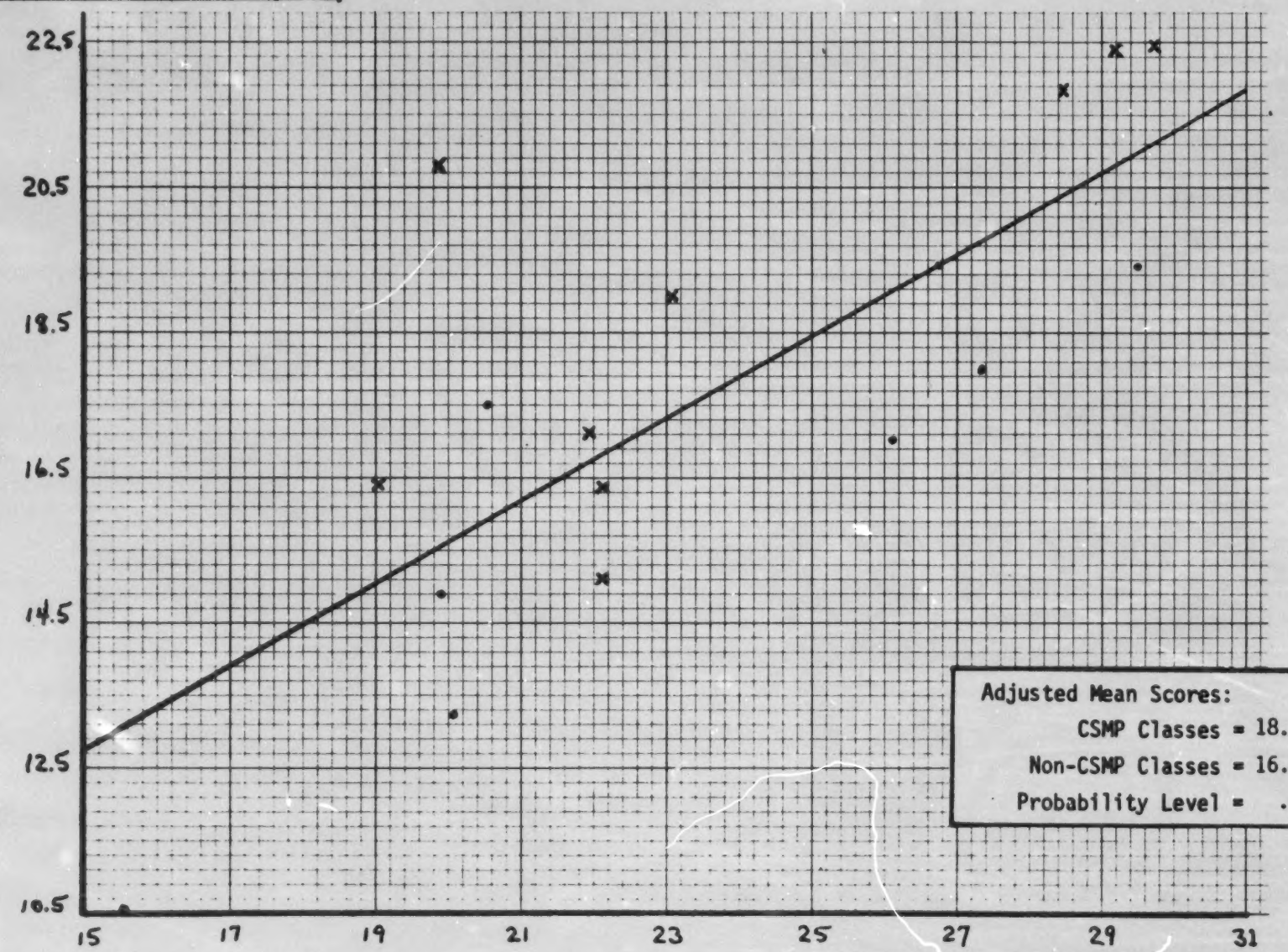
2. The items on these Mental Arithmeitc scales can be divided into two categories: those with the "box" to the right of the equal sign and those with the "box" to the left. The table below gives the results, according to these categories.

Item Category	Number of Items	Mean Percent Correct	
		CSMP	Non-CSMP
"Box" on the right	16	76	66
"Box" on the left	10	69	55

The table shows that the CSMP students tend to be relatively better on the items with the "box" on the left. Since those items have a similarity to the "reverse arrow" type of questions that are emphasized in CSMP, this relative superiority lends support to the notion that a transfer effect has occurred.

Scales C3 + C4 + C5 + C6
Mental Arithmetic (+, −, ×, ÷)

Fifth Grade Class Means
× = CSMP Class, • = Non-CSMP Class



Adjusted Mean Scores:
CSMP Classes = 18.9
Non-CSMP Classes = 16.5
Probability Level = .02

43

C7 - Standard Computation (First Alternate Form)

Test Items	Percent Correct CSMP	Non-CSMP
5739 977 59 +9686	73	56
100 -79	88	69
1362 -458	79	78
6 x 9 =	96	85
705 x219	44	36
$5\overline{)255}$	71	78
$4\overline{)27}$	70	69
$\frac{3}{8} + \frac{1}{4}$ =	11	27
$\frac{4}{5} - \frac{3}{5}$ =	62	56
$\frac{1}{3}$ of 36 =	79	40
84.2 + 3.5	75	69
0.5 - 0.5 =	66	56
0.21 x5	77	[illegible]
Number of Students	73	53
KR20 Reliability/Homogeneity	.60	.68
Mean Reading Score	21.1	17.4
Correlation of Scale and Reading Scores	.43	.69

Means by Ability Level

	1	2	3	4	Combined
CSMP ($\bar{X}_1$)	7.0	8.7	9.0	9.7	8.9
Non-CSMP ($\bar{X}_2$)	5.5	8.0	8.8	9.9	7.7
$(\bar{X}_1 - \bar{X}_2)/sd$	0.9	1.1	0.2	-0.9	2.8

Note: See page 48 and 49 for notes and graph for all three standard computation scales.

C7 - Standard Computation (Second Alternate Form)

Test Items	Percent Correct CSMP	Percent Correct Non-CSMP
$\begin{array}{r} 86 \\ +37 \\ \hline \end{array}$	96	96
$\begin{array}{r} 789 \\ 798 \\ +977 \\ \hline \end{array}$	79	70
$\begin{array}{r} 234 \\ \times 5 \\ \hline \end{array}$	85	85
$\begin{array}{r} 603 \\ \times 7 \\ \hline \end{array}$	78	72
$\begin{array}{r} 645 \\ \times 101 \\ \hline \end{array}$	54	51
$2\overline{)326}$	62	64
$12\overline{)84}$	69	66
$13\overline{)273}$	53	66
$\frac{1}{2} + \frac{1}{3} =$	18	34
$\frac{1}{3} \div \frac{1}{3} =$	29	30
$\begin{array}{r} 1.5 \\ +3.7 \\ \hline \end{array}$	84	60
$\begin{array}{r} 1.7 \\ -0.8 \\ \hline \end{array}$	56	53
$\begin{array}{r} 17.10 \\ \times 11 \\ \hline \end{array}$	37	43
Number of Students	68	53
KR20 Reliability/Homogeneity	.73	.83
Mean Reading Score	19.1	19.2
Correlation of Scale and Reading Scores	.18	.57

Means by Ability Level

	1	2	3	4	Combined
CSMP ($\bar{X}_1$)	7.2	7.8	7.8	9.4	8.0
Non-CSMP ($\bar{X}_2$)	4.4	7.3	8.8	10.0	7.8
$(\bar{X}_1 - \bar{X}_2)/sd$	2.4	0.4	-0.9	-0.5	0.0

Note: See page 48 and 49 for notes and graph for all three standard computation scales.

C7 - Standard Computation (Third Alternate Form)

Test Items	Percent Correct CSMP	Non-CSMP
$\begin{array}{r} 676 \\ +548 \\ \hline \end{array}$	91	86
$91 - 40 =$	86	86
$\begin{array}{r} 12604 \\ -8958 \\ \hline \end{array}$	45	65
$\begin{array}{r} 100 \\ \times 7 \\ \hline \end{array}$	95	86
$\begin{array}{r} 37 \\ \times 61 \\ \hline \end{array}$	61	68
$\begin{array}{r} 211 \\ \times 321 \\ \hline \end{array}$	59	53
$8\overline{)56}$	80	89
$52\overline{)260}$	64	67
$\frac{2}{6} + \frac{3}{6} =$	52	65
$\frac{4}{5} + \frac{3}{5} =$	50	70
$\frac{3}{4} \times \frac{1}{2} =$	41	46
$\begin{array}{r} 7.82 \\ -4.63 \\ \hline \end{array}$	71	75
$5.3 \div 5.3 =$	46	51
Number of Students	56	57
KR20 Reliability/Homogeneity	.77	.75
Mean Reading Score	19.2	20.6
Correlation of Scale and Reading Scores	.48	.53

Means by Ability Level

	1	2	3	4	Combined
CSMP ($\bar{X}_1$)	6 1	8.1	10.3	9.8	8.4
Non-CSMP ($\bar{X}_2$)	6.1	9.3	9.3	10.9	9.1
$(\bar{X}_1 - \bar{X}_2)/sd$	-0.2	-1.2	1.1	-2.0	-2.0

Note: See page 48 and 49 for notes and graph for all three standard computation scales.

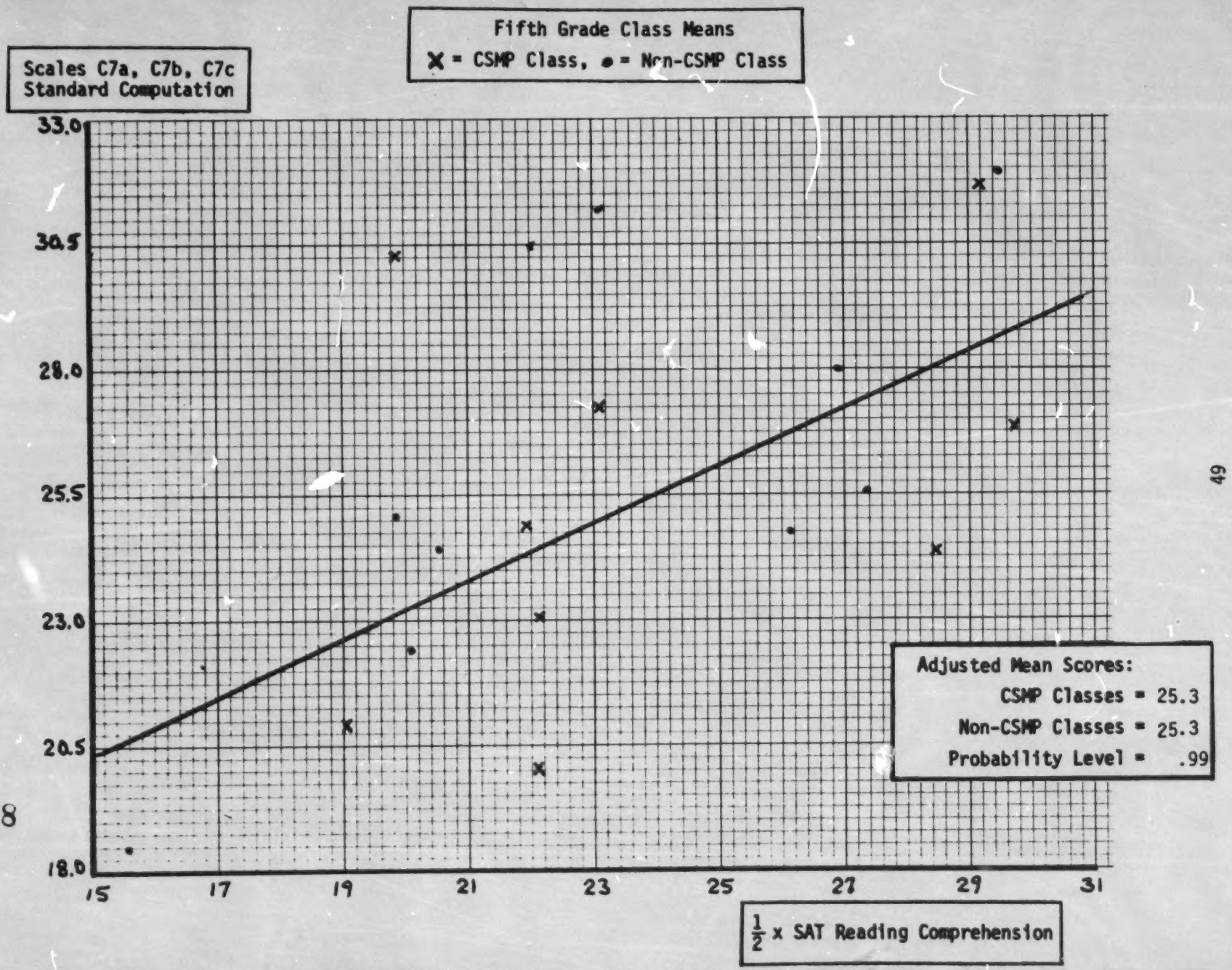
Notes on C7 (all 3 alternate forms)

It is instructive to compare the curricular groups on the mean percent correct for each of the various categories of computation items included in all forms of C7.

Item Category	No. of items	Mean Percent Correct	
		CSMP	Non-CSMP
Whole Number Addition	4	85*	77
Whole Number Subtraction	4	75	75
Whole Number Multiplication	8	72*	67
Whole Number Division	7	67	71*
Fractions: Add. and Subt.	5	39	50*
Fractions: Mult. and Div.	3	50*	39
Decimals: Add. and Subt.	5	70*	63
Decimals: Mult. and Div.	3	53*	48

Note that CSMP students were better in five of the categories, and Non-CSMP students in two. Those latter two categories (Whole Dumber Division, and the Addition and Subtraction of Fractions) are usually the two given heaviest concentration in the latter part of a traditional fourth grade.

Note also that the CSMP advantage in Whole Number Addition and Multiplication is consistent with results in other years and grades. The CSMP superiority in the other three categories probably reflects CSMP much earlier emphasis on these kind of fractions and on decimals.


Fifth Grade Class Means
X = CSMP Class, • = Non-CSMP Class
Scales C7a, C7b, C7c
Standard Computation
33.0
30.5
28.0
25.5
23.0
20.5
18.0
15
17
19
21
23
25
27
29
31
½ x SAT Reading Comprehension
Adjusted Mean Scores:
CSMP Classes = 25.3
Non-CSMP Classes = 25.3
Probability Level = .99

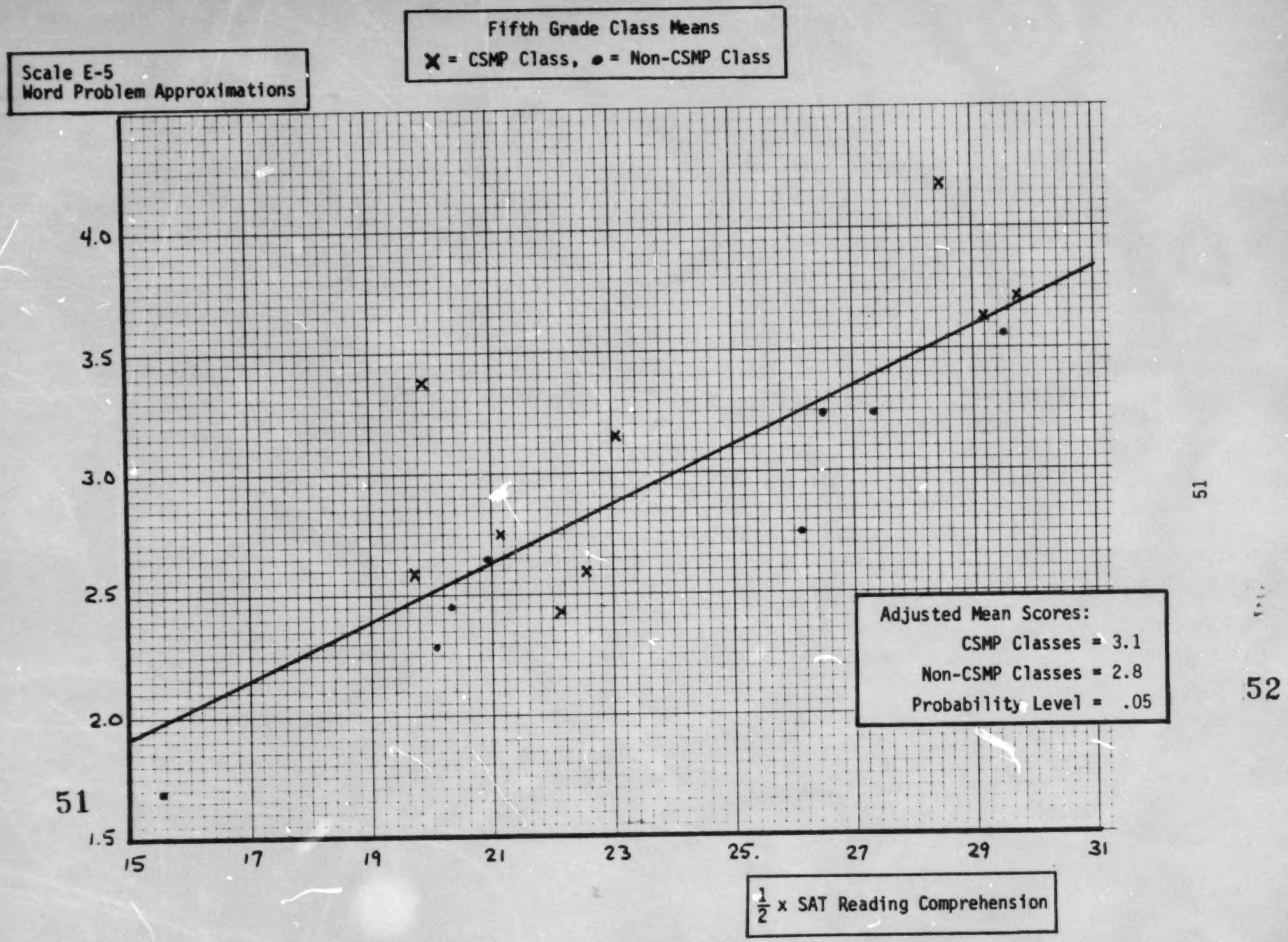
49

E5 - Word Problem Approximations

Test Items	Percent Correct CSMP	Percent Correct Non-CSMP	Bi-Serial C/N-
1. George has $902. He wants to buy a piano costing $1,348. About how much more money does he need? $250 $350 $450 $550	58	36	3,2
2. Tom and Bill had a long distance race. Bill ran it in 32 minutes, 2 seconds. Tom ran it in 29 minutes, 34 seconds. About how much did Tom win by? $1\frac{1}{2}$ minutes $2\frac{1}{2}$ minutes $3\frac{1}{2}$ minutes $4\frac{1}{2}$ minutes	20	28	1, $\hat{0}$
3. A white pine will grow to be about 100 feet tall. A blue spruce will grow to be about 30 feet tall. A white pine will grow about how many times as tall as a blue spruce? 2 times 3 times 5 times 10 times	68	53	3,3
4. Martha can walk 2 blocks in 5 minutes. About how many blocks can she walk in 13 minutes? 5 blocks 10 blocks 15 blocks 18 blocks	52	44	3,3
5. Susan has $131. Chairs cost $32. About how many chairs can Susan buy? 2 chairs 4 chairs 6 chairs 10 chairs	72	60	2,4
6. It takes 3 or 4 cups of water to cook one dozen eggs. Mary has to cook 5 dozen eggs. About how many cups of water will she need? 8 cups 12 cups 16 cups 25 cups	47	48	3,3
Number of Students	192	160	
KR20 Reliability/Homogeneity	.47	.47	
Mean Reading Score	19.9	19.2	
Correlation of Scale and Reading Scores	.31	.50	

Means by Ability Level

	1	2	3	4	Combined
CSMP ($\bar{X}_1$)	2.6	3.0	3.0	4.0	3.2
Non-CSMP ($\bar{X}_2$)	1.4	2.6	2.9	3.8	2.7
$(\bar{X}_1-\bar{X}_2)/sd$	3.8	1.7	0.6	0.5	2.9

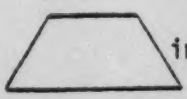
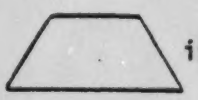
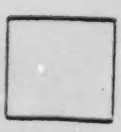

Fifth Grade Class Means
× = CSMP Class, • = Non-CSMP Class
Scale E-5
Word Problem Approximations
4.0
3.5
3.0
2.5
2.0
1.5
15
17
19
21
23
25.
27
29
31
Adjusted Mean Scores:
CSMP Classes = 3.1
Non-CSMP Classes = 2.8
Probability Level = .05
½ x SAT Reading Comprehension

51

51

52

G1 - Geometric Congruencies

Test Items (See Appendix B for fuller description)	Percent Correct CSMP	Percent Correct Non-CSMP
in 2 congruent parts*	97	97
in 3 congruent parts	85	85
in 4 congruent parts	83	81
in 6 congruent parts	63	40
 in 2 congruent parts	78	76
 in 3 congruent parts	44	28
in 4 congruent parts	65	64
 in 6 congruent parts	66	68
Number of Students	197	165
KR20 Reliability/Homogeneity	.76	.69
Mean Reading Score	19.8	19.1
Correlation of Scale and Reading Scores	.38	.41

*Note: The word "congruent" was not used with students; for each given shape the parts had to be given the "same size and shape"

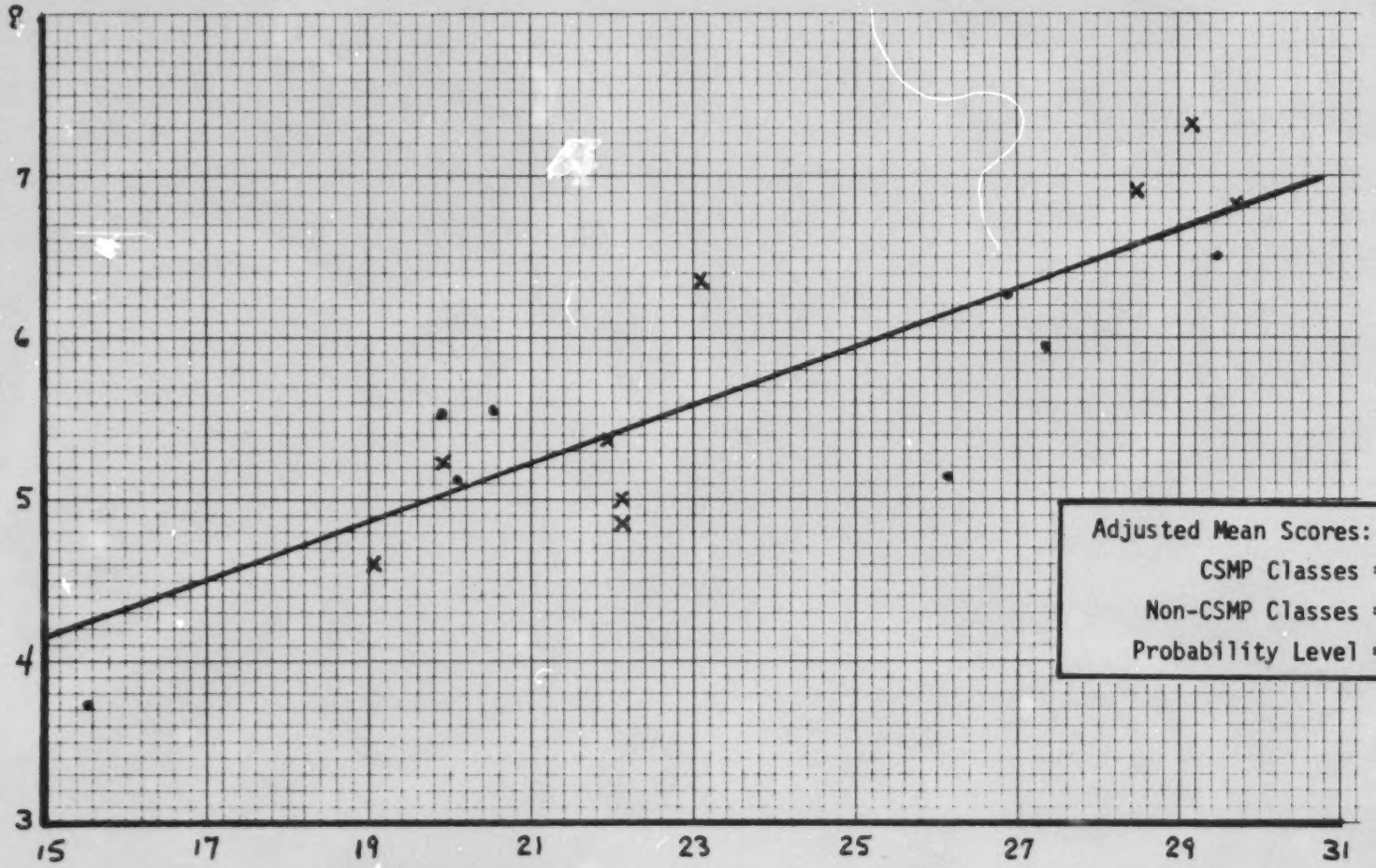
Means by Ability Level

	1	2	3	4	Combined
CSMP ($\bar{X}_1$)	4.9	4.9	6.1	7.2	5.8
Non-CSMP ($\bar{X}_2$)	4.3	5.3	5.6	6.4	5.4
$(\bar{X}_1-\bar{X}_2)/sd$	1.2	-0.8	1.4	2.4	1.7

Fifth Grade Class Means

× = CSMP Class, • = Non-CSMP Class

Scale G1 -
Geometric Congruencies



Adjusted Mean Scores:

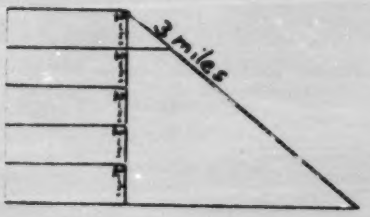
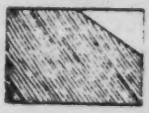
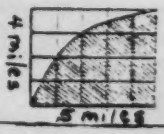
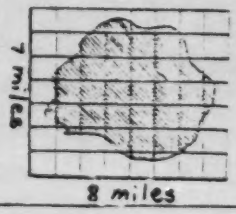
CSMP Classes = 5.8

Non-CSMP Classes = 5.5

Probability Level = .36

$\frac{1}{2}$ x SAT Reading Comprehension

G2 - Geometric Approximations: 1D

Test Items	Percent Correct CSMP	Non-CSMP
The double line is about _____ miles	84	76
The double line is about _____ miles	9	13
 The double line is about _____ miles	17	13
← The white part is 25 square miles. ← The shaded part is about _____ square miles.	57	39
 ← The white part is 100 square miles. ← The shaded part is about _____ square miles.	14	7
 The shaded part is about _____ square miles.	34	28
 The shaded part is about _____ square miles.	14	18
Number of Students	192	160
KR20 Reliability/Homogeneity	.44	.49
Mean Reading Score	19.9	19.2
Correlation of Scale and Reading Scores	.41	.48

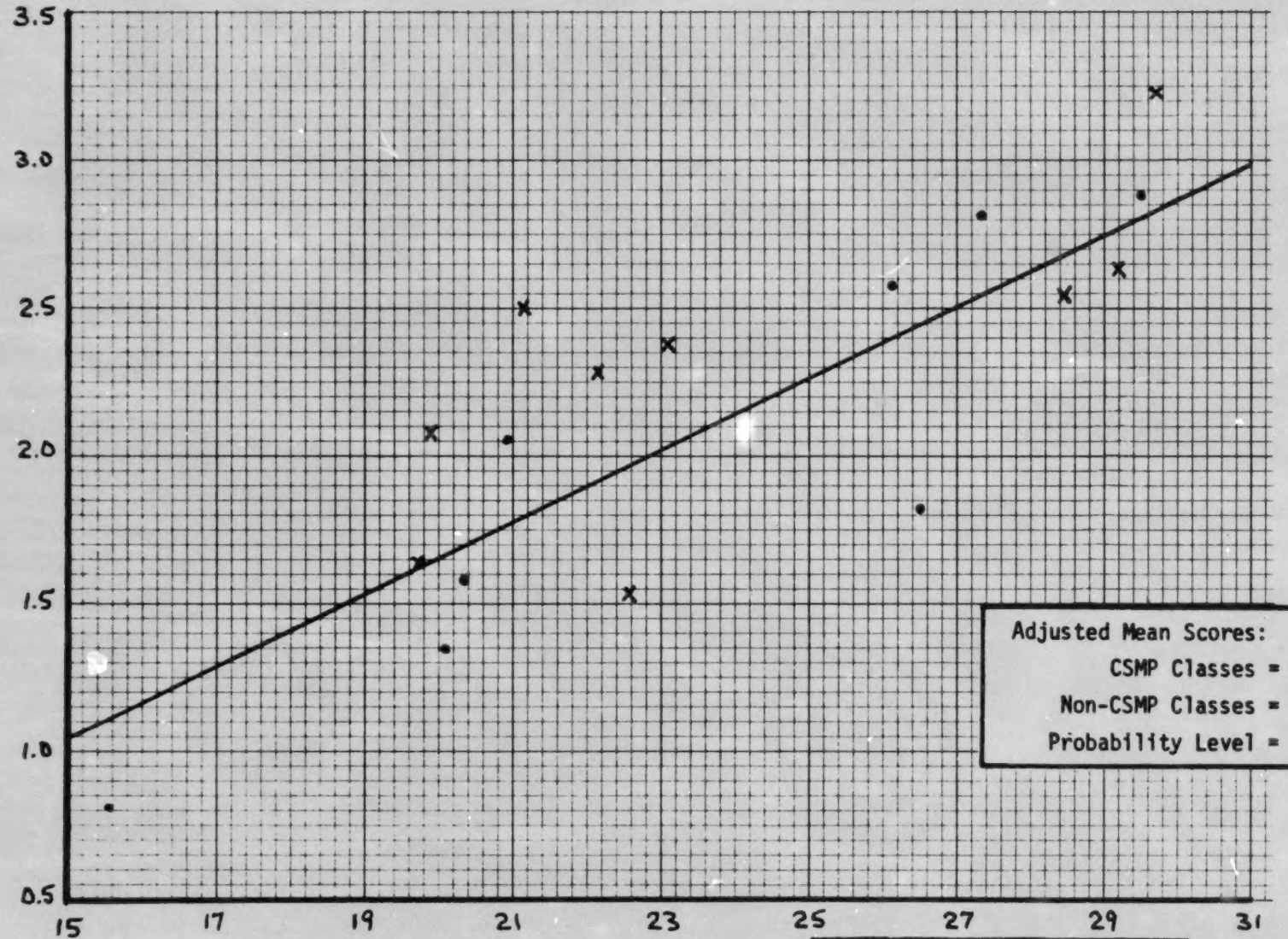
Means by Ability Level

	1	2	3	4	Combined
CSMP ($\bar{X}_1$)	1.3	2.3	2.3	3.0	2.3
Non-CSMP ($\bar{X}_2$)	1.0	1.7	2.4	2.6	1.9
$(\bar{X}_1 - \bar{X}_2)/sd$	1.0	2.7	-0.3	1.0	2.2

Fifth Grade Class Means

× = CSMP Class, • = Non-CSMP Class

G2 -
Geometric Approximations



Adjusted Mean Scores:

CSMP Classes = 2.3

Non-CSMP Classes = 2.0

Probability Level = .19

57

55

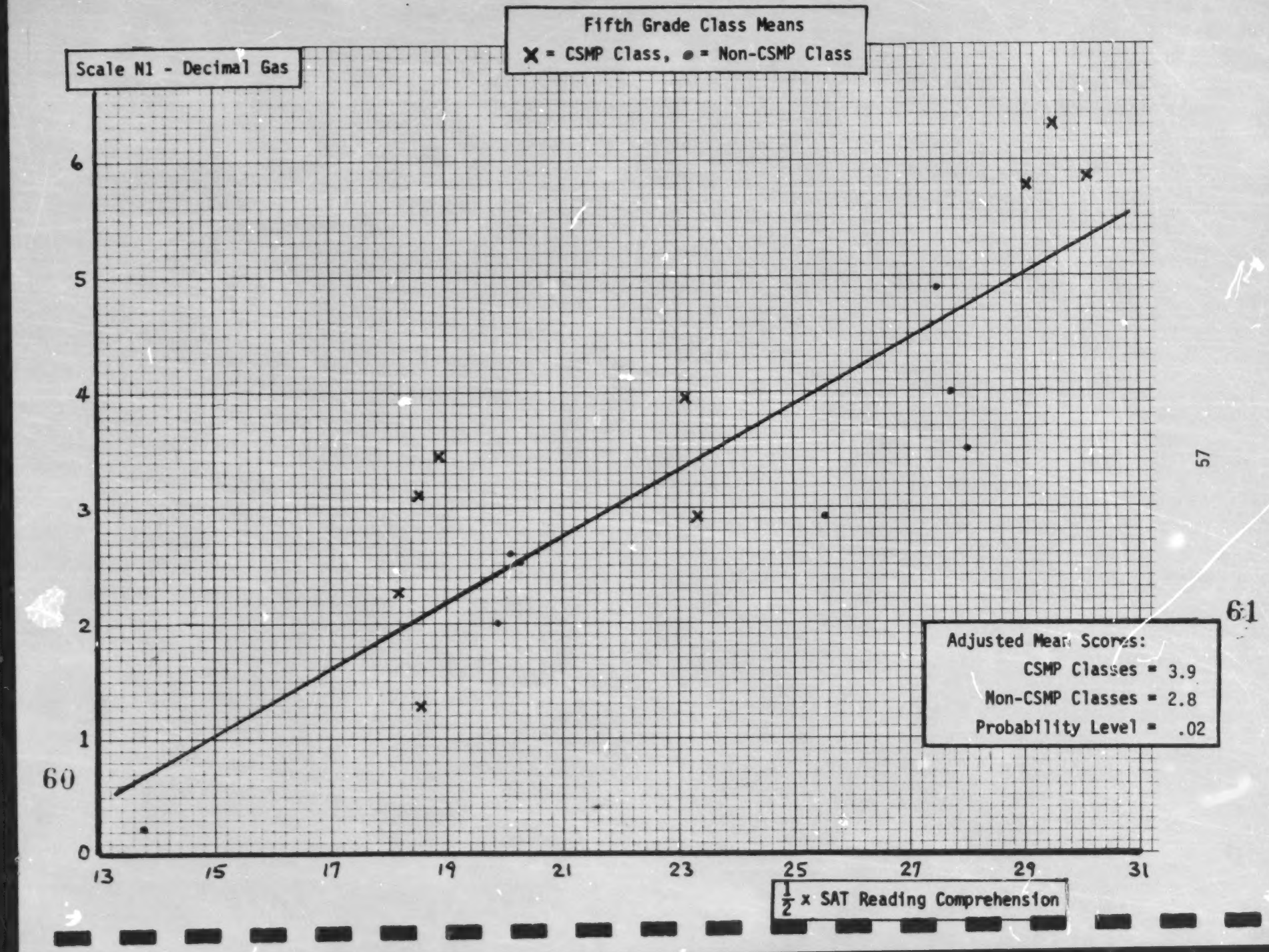
58

N1 - Decimal Gas

Test Items	Percent Correct CSMP	Non-CSMP
Buying Gas Six brothers were buying gas. Each brother bought 6.5 gallons.		
1. John has 6.5 gallons. He uses up four gallons. How much gas will he have left? _____	74	62
2. Bill has 6.5 gallons. He buys another half gallon. How much gas will he have then? _____	61	38
3. Joe has 6.5 gallons. He sells each gallon for $2. How much money will he get altogether? _____	36	31
4. Tom has 6.5 gallons. He buys 3.5 more gallons. How much gas will he have then? _____	79	60
5. Ron has 6.5 gallons. Next week he will use ten times this much. How much gas will he use next week? _____	47	25
6. Ken has 6.5 gallons of gas. He gives away half of it. How much gas will he have left? _____	39	6
7. Mrs. Jones had this much gas. Each can holds 1 gallon. How many gallons did she have? Circle the best answer. 3 gallons 3.3 gallons $3\frac{1}{2}$ gallons 3.8 gallons 4 gallons	55	48
Number of Students	98	84
KR20 Reliability/Homogeneity	.83	.78
Mean Reading Score	19.6	18.7
Correlation of Scale and Reading Scores	.62	.53

Means by Ability Level

	1	2	3	4	Combined
CSMP ($\bar{X}_1$)	1.9	3.1	4.6	5.7	3.9
Non-CSMP ($\bar{X}_2$)	0.9	2.7	3.9	4.2	2.7
$(\bar{X}_1-\bar{X}_2)/sd$	1.9	0.8	2.5	2.7	3.9


Fifth Grade Class Means
× = CSMP Class, • = Non-CSMP Class
Scale N1 - Decimal Gas
0
1
2
3
4
5
6
13
15
17
19
21
23
25
27
29
31
½ x SAT Reading Comprehension
Adjusted Mean Scores:
CSMP Classes = 3.9
Non-CSMP Classes = 2.8
Probability Level = .02

60

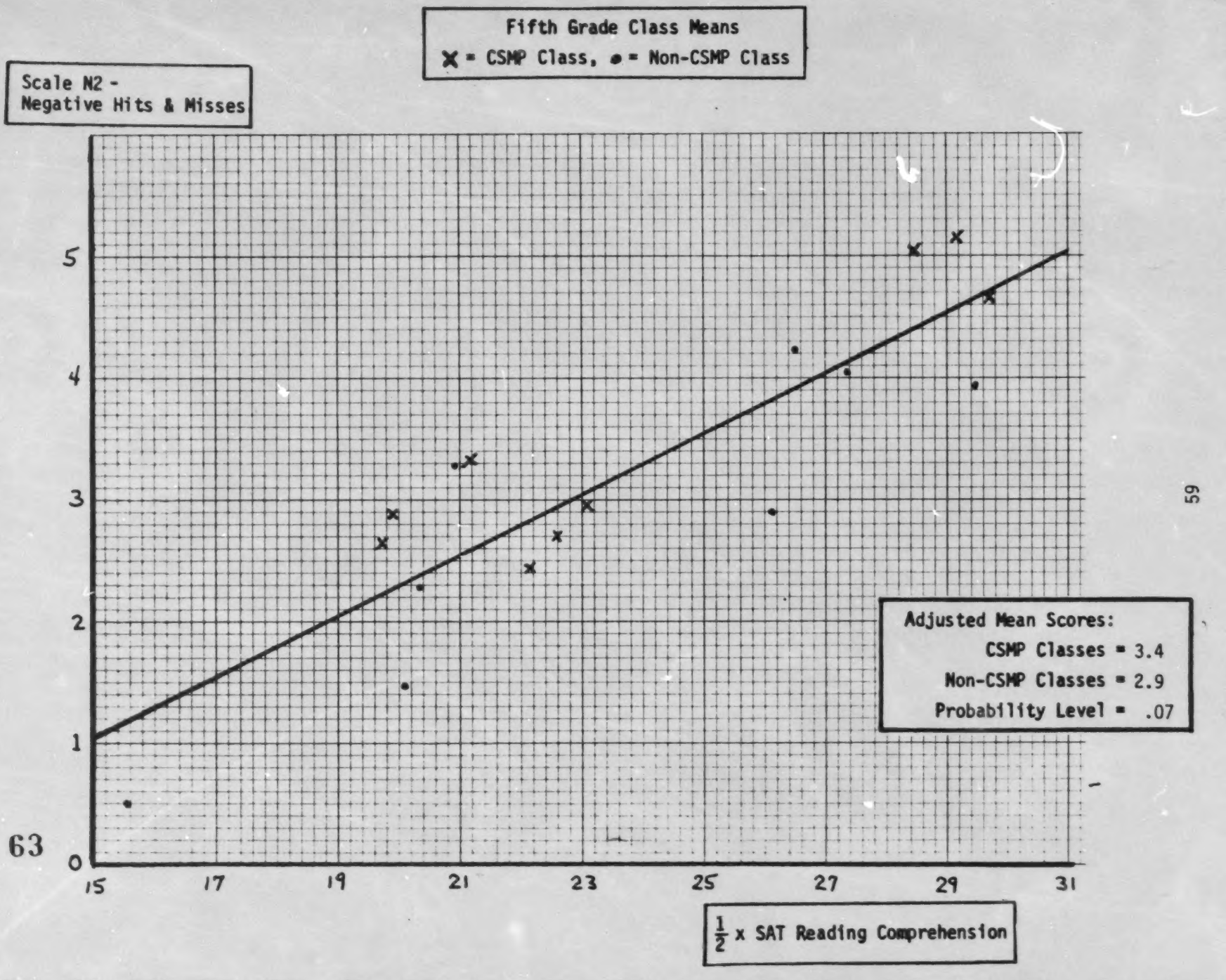
61

N2 - Negative Hits and Misses

Test Items	Percent Correct: CSMP	Percent Correct: Non-CSMP
Bill — Started with a score of: 0; Number of Hits: 2; Number of Misses: 3; Ended with a score of:	73	62
Ellen — Started with a score of: 3 below zero; Number of Hits: 2; Number of Misses: 0; Ended with a score of:	61	51
May — Started with a score of: ; Number of Hits: 1; Number of Misses: 6; Ended with a score of: 3	50	38
Jan — Started with a score of: 3; Number of Hits: 0; Number of Misses: 7; Ended with a score of:	65	55
John — Started with a score of: ; Number of Hits: 2; Number of Misses: 0; Ended with a score of: 15 below zero	46	32
Peter — Started with a score of: 10 below zero; Number of Hits: 1; Number of Misses: ; Ended with a score of: 12 below zero	49	34
Number of Students	192	160
KR20 Reliability/Homogeneity	.86	.85
Mean Reading Score	19.9	19.2
Correlation of Scale and Reading Scores	.53	.57

Means by Ability Level

	1	2	3	4	Combined
CSMP ($\bar{X}_1$)	1.8	2.7	4.0	5.2	3.5
Non-CSMP ($\bar{X}_2$)	0.9	2.3	3.3	4.3	2.7
$(\bar{X}_1-\bar{X}_2)/sd$	2.6	1.0	1.5	2.2	3.1


Fifth Grade Class Means
× = CSMP Class, • = Non-CSMP Class
Scale N2 -
Negative Hits & Misses
0
1
2
3
4
5
15
17
19
21
23
25
27
29
31
½ x SAT Reading Comprehension
Adjusted Mean Scores:
CSMP Classes = 3.4
Non-CSMP Classes = 2.9
Probability Level = .07

P1 - Spinners

Test Items (See Appendix B for full description)	Allowable Range	Percent Correct CSMP	Percent Correct Non-CSMP
SUPPOSE WE PLAY THE GAME 100 TIMES			
About how many times would we get a score of 1?____	(40-60)	51	44
About how many times would we get a score of 2?____	(20-30)	61	61
About how many times would we get a score of 3?____	(10-20)	52	52
About how many times would we get a score of 4?____	(10-20)	51	50
About how many times would we get a score of less than 4? ____	(80-90)	32	25
First 4 responses add to 100		42	39
Third response = Fourth response		41	30
First 3 responses = Fifth response		32	25
First response > second > third & fourth		40	27
Last 2 responses add to 100		31	27
SCORES YOU COULD GET 25, 17, 35, 15, 27, 37			
SUPPOSE WE PLAY THE GAME 100 TIMES			
About how many times would we get a score of 15? _____	(20-30)	29	32
About how many times would we get a score of 17? ____	(20-30)	31	39
About how many times would we get a score of 25? ____	(10-25)	43	54
About how many times would we get a score of less than 20?____	(40-60)	17	11
About how many times would we get a score of less than 30? ____	(60-90)	26	15
First response = Second response		45	43
First & second responses > third response		62	61
First + second response = fourth response		14	12
Fifth response = fourth response		58	59
Number of Students		197	165
KR20 Reliability/Homogeneity		.85	.83
Mean Reading Score		19.8	19.1
Correlation of Scale and Reading Scores		.41	.53

Means by Ability Level

	1	2	3	4	Combined
CSMP ($\bar{X}_1$)	4.7	6.0	8.4	11.0	7.6
Non-CSMP ($\bar{X}_2$)	4.2	6.1	7.2	10.6	7.1
$(\bar{X}_1-\bar{X}_2)/sd$	0.6	-0.1	1.2	0.5	0.6

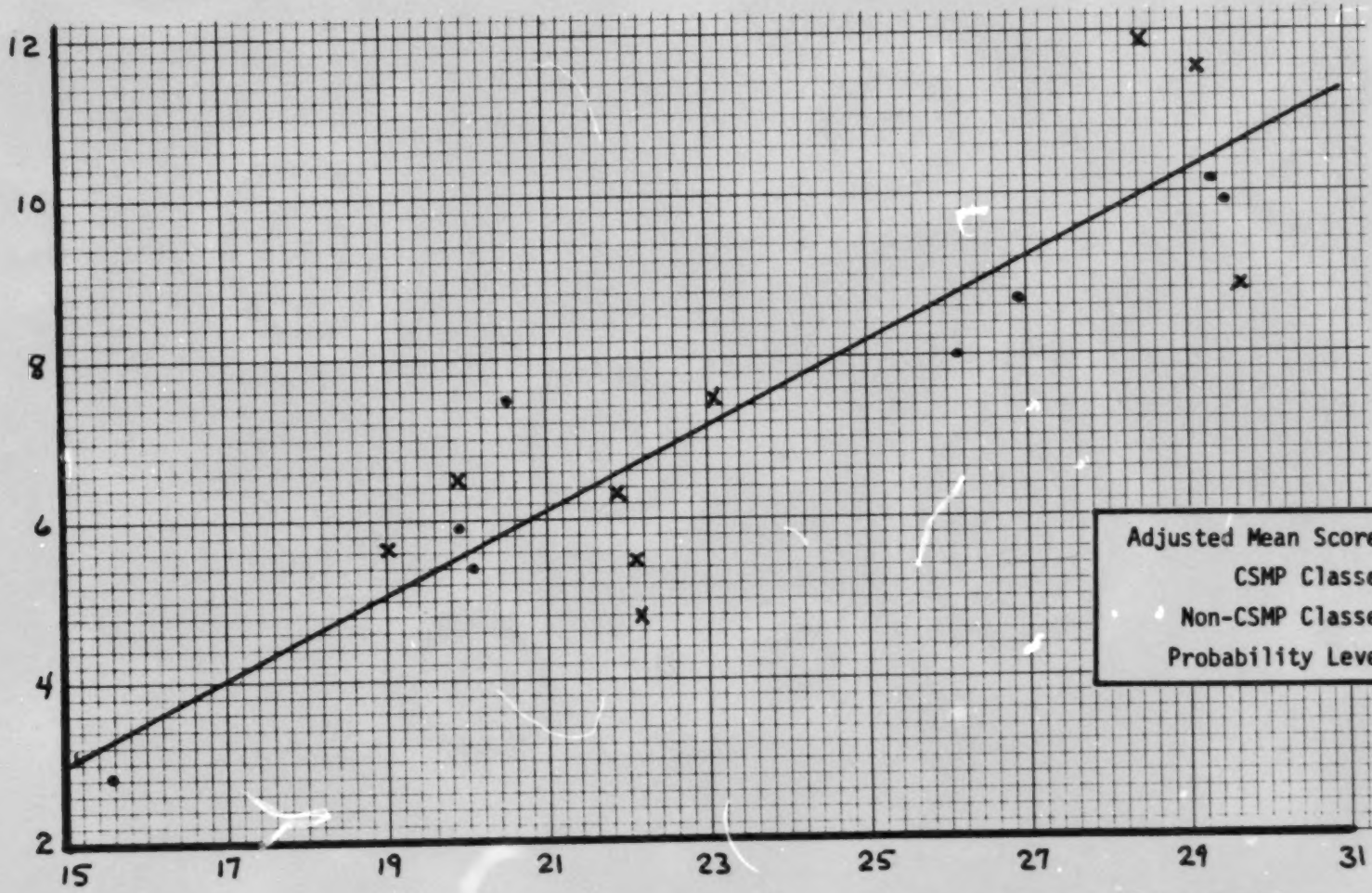
Note: [1]Only the first five items on each section appeared in the test booklet. The coding of the other nine items was based on responses to these first sets of items.

[2]Had the scoring of the first five items in each set required the exact answer (rather than a range of "close" answers), the mean scores would have been 5.6 for CSMP and 4.4 for Non-CSMP (versus 7.6 and 7.1 for the scoring method actually used).

Fifth Grade Class Means

× = CSMP Class, • = Non-CSMP Class

Scale P1 - Spinners



Adjusted Mean Scores:

CSMP Classes = 7.4

Non-CSMP Classes = 7.5

Probability Level = .94

$\frac{1}{2}$ x SAT Reading Comprehension

P2 - Which is More Likely

Test Items			Percent Correct CSMP	Percent Correct Non-CSMP
	Which is More Likely this? ↓	or this? ↓ or both the same?		
Ⓐ Flip 3 coins at once	three heads	① three tails	41	43
	three heads	② two heads and one tail	73	76
	one head and two tails	③ two heads and one tail	41	36
Ⓑ Roll a pair of dice	total of 2	① total of 12	27	25
	total of 7	② total of 11	55	58
	total of 4 or 5	③ total of 7 or 8	48	48
	total of less than 5	④ total of more than 10	53	53
Ⓒ Spin this spinner	score of 1	① score of 2	72	71
	score of 2	② score of 3 or 4	23	24
	score of less than 2	③ score of more than 2	54	50
	score of 1	④ score of 2 or 3	60	52
		Number of Students	192	160
		KR20 Reliability/Homogeneity	.64	.66
		Mean Reading Score	19.9	19.2
		Correlation of Scale and Reading Scores	.48	.40

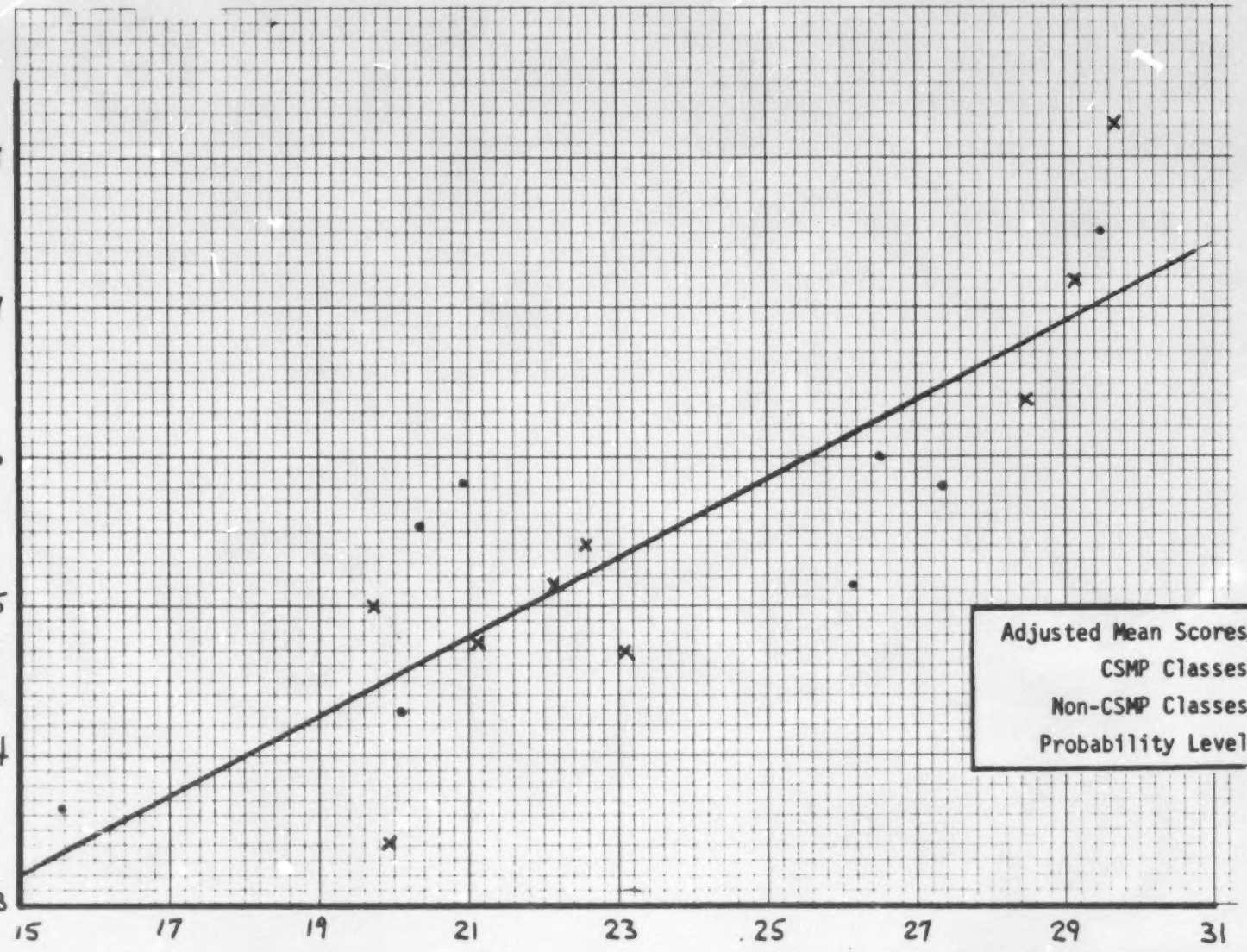
Means by Ability Level

	1	2	3	4	Combined
CSMP ($\bar{X}_1$)	3.9	4.6	5.7	7.4	5.5
Non-CSMP ($\bar{X}_2$)	3.9	5.0	5.7	6.8	5.4
$(\bar{X}_1-\bar{X}_2)/sd$	0.1	-0.7	0.1	0.5	-0.1

Fifth Grade Class Means

× = CSMP Class, • = Non-CSMP Class

Scale P2 - Which is More Likely



Adjusted Mean Scores:

CSMP Classes = 5.5

Non-CSMP Classes = 5.6

Probability Level = .84

$\frac{1}{2}$ x SAT Reading Comprehension

69

63

70

Test Items (See Appendix B for fuller description)	Percent Correct CSMP	Non-CSMP

Cost of Pipe

Width \ Length	100'	300'	600'	1,000'
4"	$50	$150	$300	$500
8"	$70	$210	$420	$700
12"	$90	$270	$540	$900
16"	$110	$330	$660	$1100

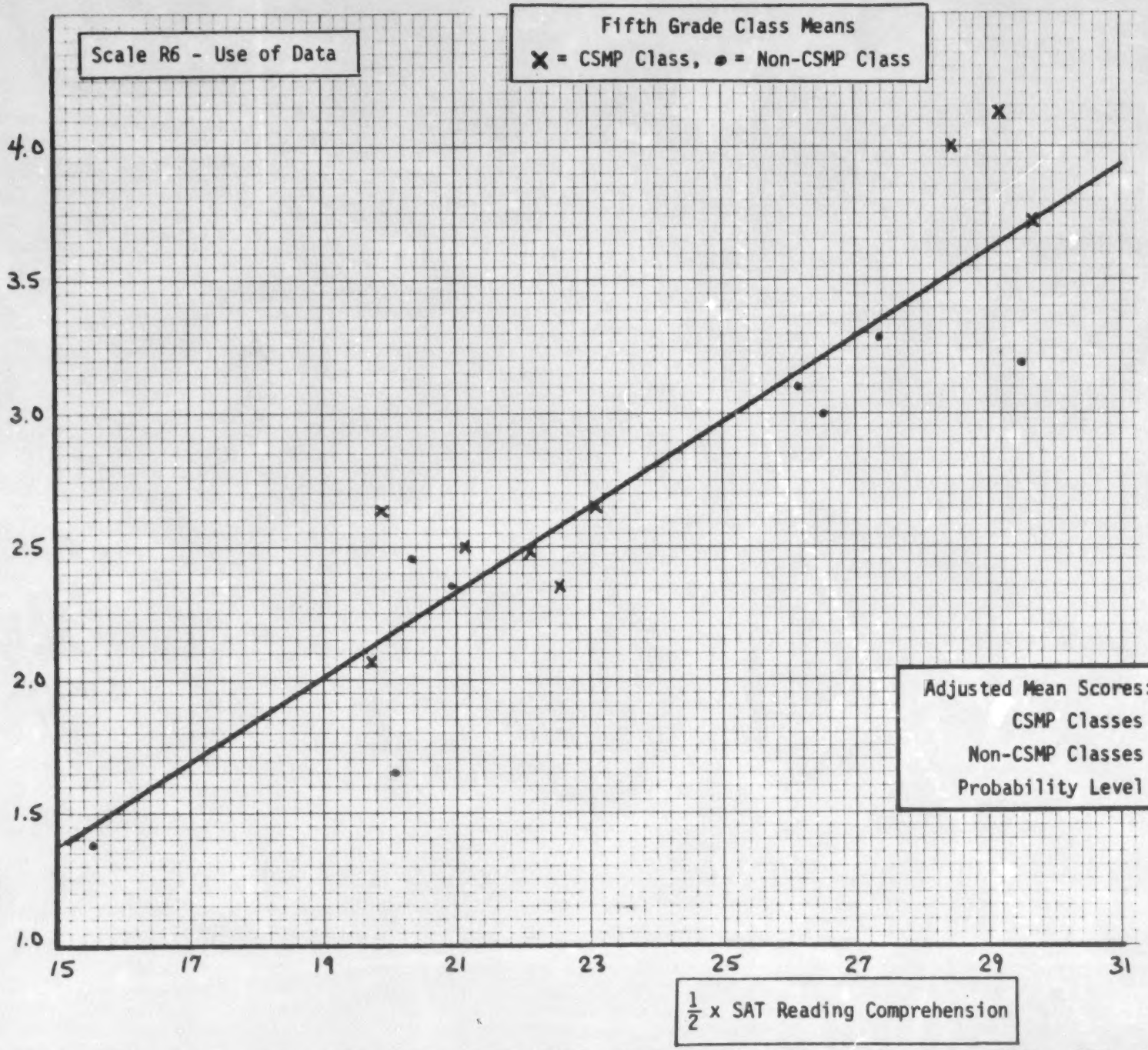
HOW MUCH DOES IT COST TO BUY PIPE WHICH IS:

4" by 200'	$_____	75	69
5" by 100'	$_____	21	10
10" by 1000'	$_____	76	66
4" by 500'	$_____	35	31
4" by 800'	$_____	55	56
20" by 1000'	$_____	20	15
6" by 200'	$_____	9	6

Number of Students	192	160
KR20 Reliability/Homogeneity	.62	.69
Mean Reading Score	19.9	19.2
Correlation of Scale and Reading Scores	.49	.51

Means by Ability Level

	1	2	3	4	Combined
CSMP ($\bar{X}_1$)	1.7	2.6	3.1	4.1	2.9
Non-CSMP ($\bar{X}_2$)	1.2	2.4	3.0	3.4	2.5
$(\bar{X}_1-\bar{X}_2)/sd$	1.2	0.5	0.5	1.9	1.9

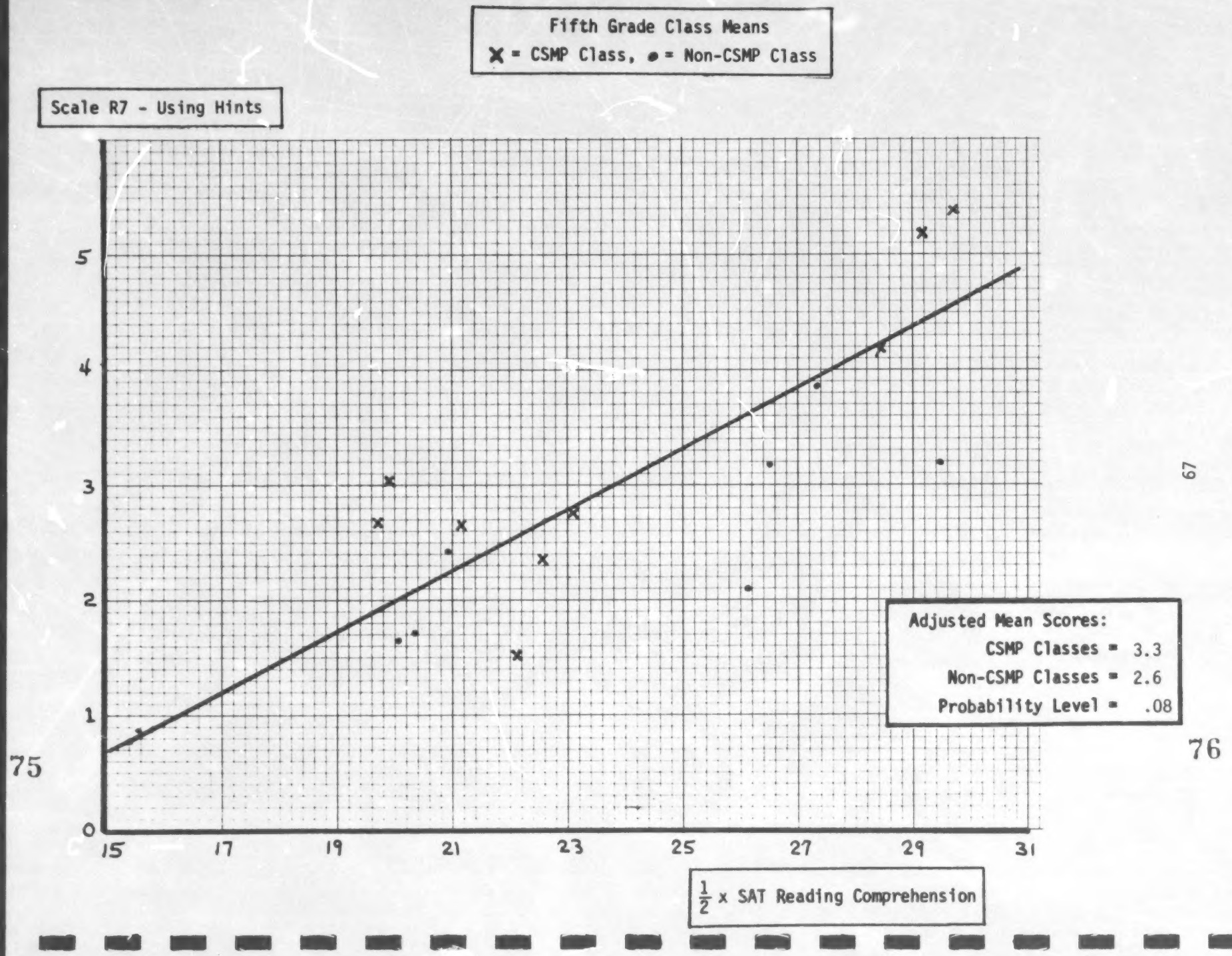

Scale R6 - Use of Data
Fifth Grade Class Means
× = CSMP Class, • = Non-CSMP Class
4.0
3.5
3.0
2.5
2.0
1.5
1.0
15
17
19
21
23
25
27
29
31
½ x SAT Reading Comprehension
Adjusted Mean Scores:
CSMP Classes = 2.9
Non-CSMP Classes = 2.6
Probability Level = .06

R7 - Using Hints

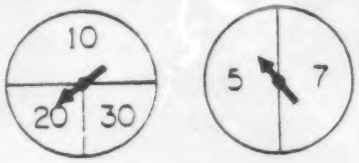
Test Items (in abbreviated format)	Percent Correct CSMP	Percent Correct Non-CSMP
Addition		
$91.17 + 83.22 + 47.65 + 65.89 + 24.74 + 72.56 = $385.23	47	37
$89.87 + $121.92 = $211.92	18	13
$44.25 + $374.42 = $418.67	34	15
$68.89 + $362.09 = $430.98	39	26
$14.98 + $378.31 = $393.29	18	12
$57.89 + $374.33 = $432.22	27	14
$47.00 + $386.00 = $433.00	21	14
Multiplication		
4 × 74 = 444	47	33
22 × 22 = 484	36	32
16 × 14 = 224	30	34
Number of Students	192	160
KR20 Reliability/Homogeneity	.83	.74
Mean Reading Score	19.9	19.2
Correlation of Scale and Reading Scores	.46	.44

Means by Ability Level

	1	2	3	4	Combined
CSMP ($\bar{X}_1$)	1.6	2.2	3.5	5.5	3.3
Non-CSMP ($\bar{X}_2$)	1.1	2.1	2.3	3.6	2.3
$(\bar{X}_1 - \bar{X}_2)/sd$	1.0	0.4	2.4	2.4	3.3


Fifth Grade Class Means
× = CSMP Class, • = Non-CSMP Class
Scale R7 - Using Hints
5
4
3
2
1
0
15
17
19
21
23
25
27
29
31
Adjusted Mean Scores:
CSMP Classes = 3.3
Non-CSMP Classes = 2.6
Probability Level = .08
½ x SAT Reading Comprehension

U1 - Elucidation

Test Items	Number of other possible correct answers	Mean Number of Correct answers		Mean Number of incorrect responses	
		CSMP	Non-CSMP	CSMP	Non-CSMP
① Spinner game Your score is the total from the two spinners  Possible scores 25.	5	4.5	4.0	0.5	1.0
② Start at zero Get to 24 on the dice. Counting by 2.	7	4.1	3.9	0.9	1.5
③ Each hit, gain 5 points Each miss, lose 1 point Take 5 throws Possible scores 25.	5	2.4	1.6	2.3	3.2
④ Multiple of 2 Multiple of 3 Smaller than 50 All true for these numbers 24.	7	3.9	2.8	1.2	1.8

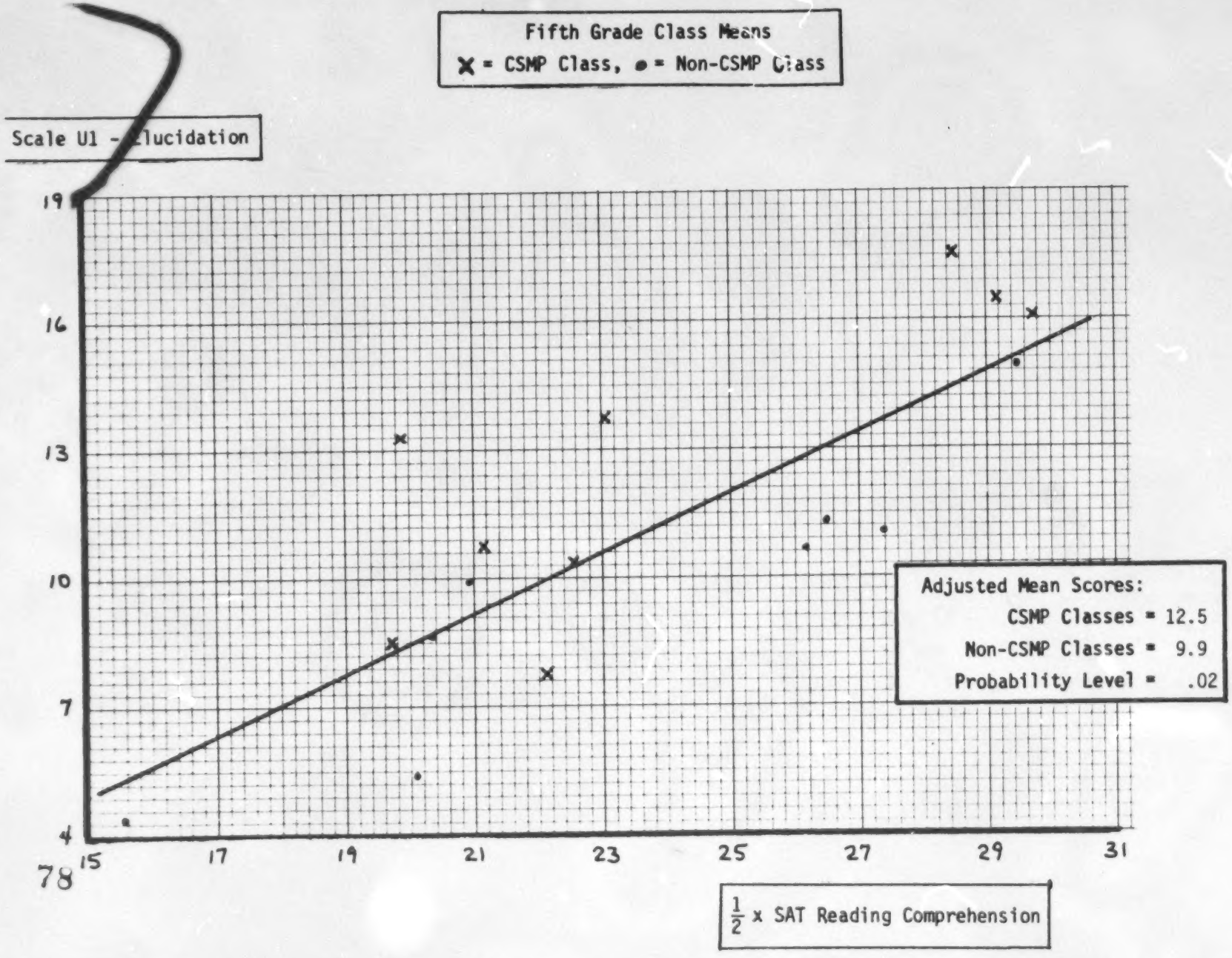
	CSMP	Non-CSMP
Number of Students	192	160
Mean Reading Score	19.9	19.2
Correlation of Scale and Reading Scores	.46	.55

Means by Ability Level

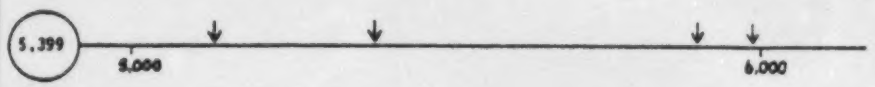
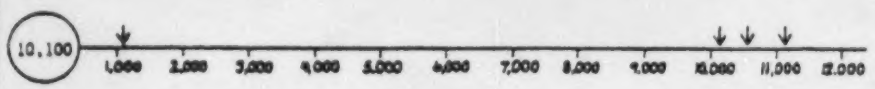
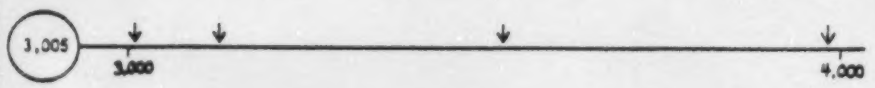
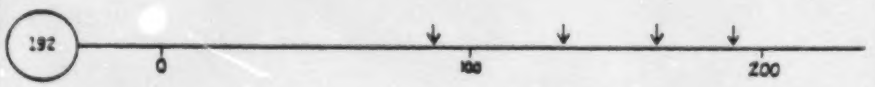
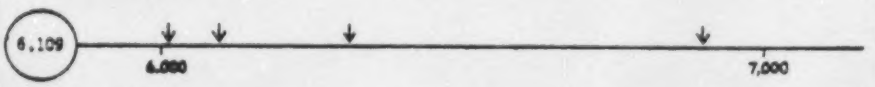
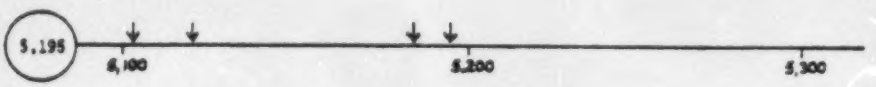
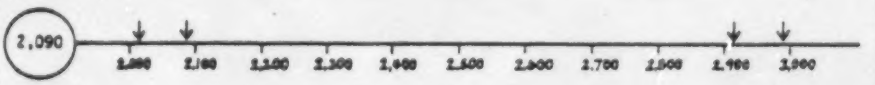
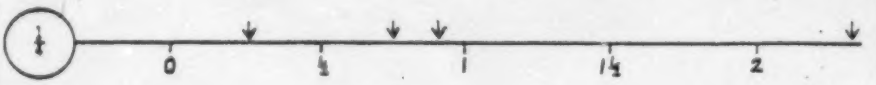
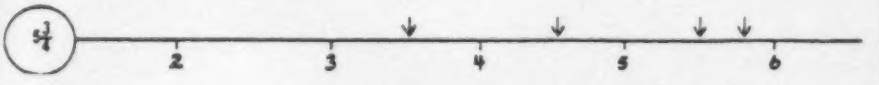
	1	2	3	4	Combined
CSMP ($\bar{X}_1$)	8.9	10.1	13.0	18.3	12.7
Non-CSMP ($\bar{X}_2$)	3.8	9.6	9.9	13.8	9.3
$(\bar{X}_1-\bar{X}_2)/sd$	2.8	0.4	2.6	3.1	4.9

Notes:

1. For the purposes of obtaining total scores on this scale, each item was scored as (number of correct responses) minus (number of incorrect responses) unless that number was less than zero, in which case the score was zero. Hence a student could score as high as 24 on this scale.
2. The percentages of students getting all possible correct answers was, for items 1-4 respectively: 79% CSMP versus 70% Non-CSMP, 15% versus 5%, 27% versus 12%, and 28% versus 12%.


Fifth Grade Class Means
× = CSMP Class, • = Non-CSMP Class
Scale U1 - Elucidation
19
16
13
10
7
4
15
17
19
21
23
25
27
29
31
Adjusted Mean Scores:
CSMP Classes = 12.5
Non-CSMP Classes = 9.9
Probability Level = .02
1/2 x SAT Reading Comprehension

V1 - Arrowed Number Lines

Test Items	Percent Correct CSMP	Percent Correct Non-CSMP
	76	60
	54	55
	51	51
	55	53
	45	43
	27	20
	69	68
	81	70
	42	41
	60	52
	37	39
Number of Students	192	160
KR20 Reliability/Homogeneity	.79	.78
Mean Reading Score	19.9	19.2
Correlation of Scale and Reading Scores	.56	.59

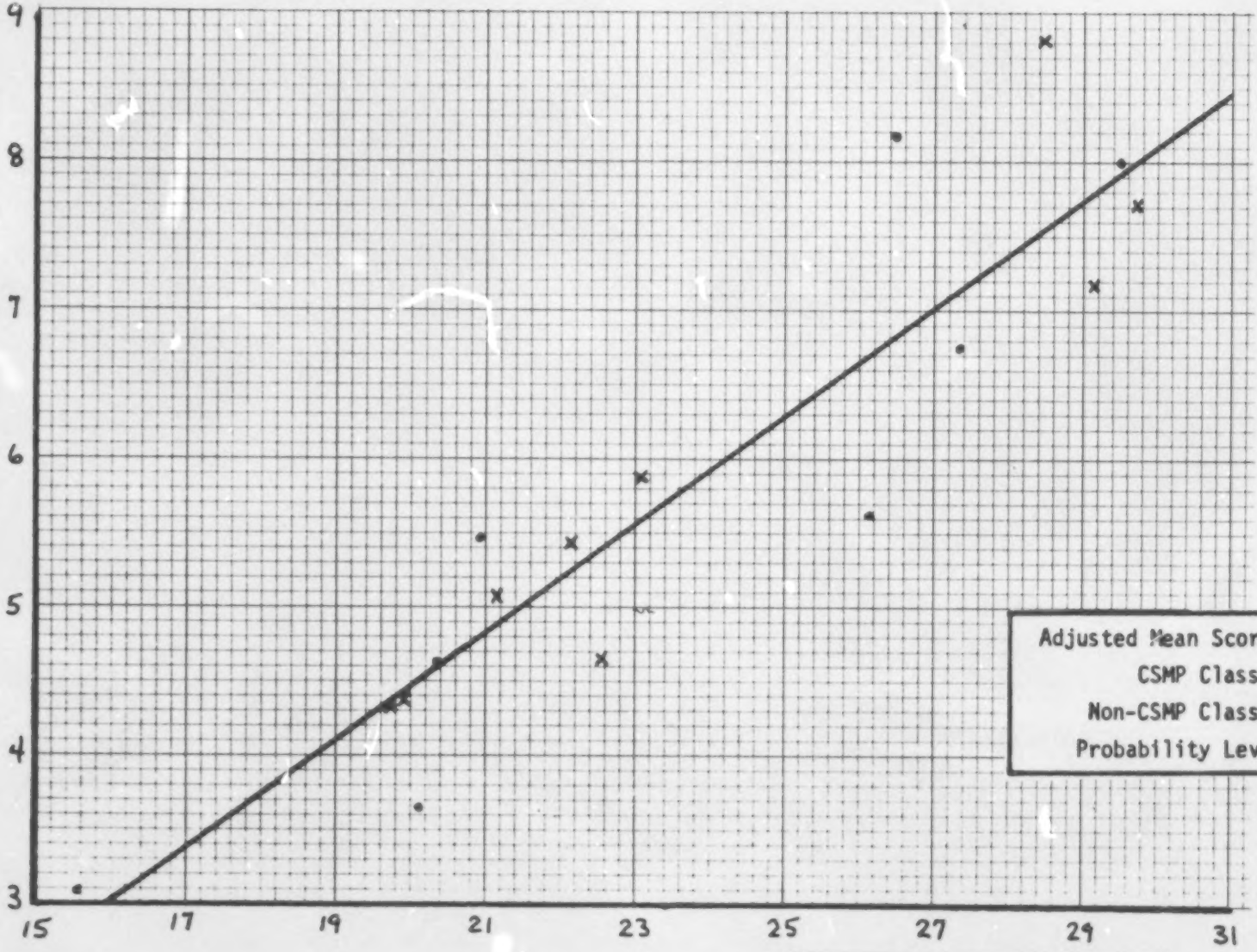
Means by Ability Level

	1	2	3	4	Combined
CSMP ($\bar{X}_1$)	4.2	4.5	6.4	8.6	6.0
Non-CSMP ($\bar{X}_2$)	2.8	5.4	5.9	7.7	5.5
$(\bar{X}_1-\bar{X}_2)/sd$	2.6	-1.7	0.9	0.9	1.1

Fifth Grade Class Means

× = CSMP Class, • = Non-CSMP Class

V1 - Place Value



Adjusted Mean Scores:

CSMP Classes = 5.9

Non-CSMP Classes = 5.8

Probability Level = .85

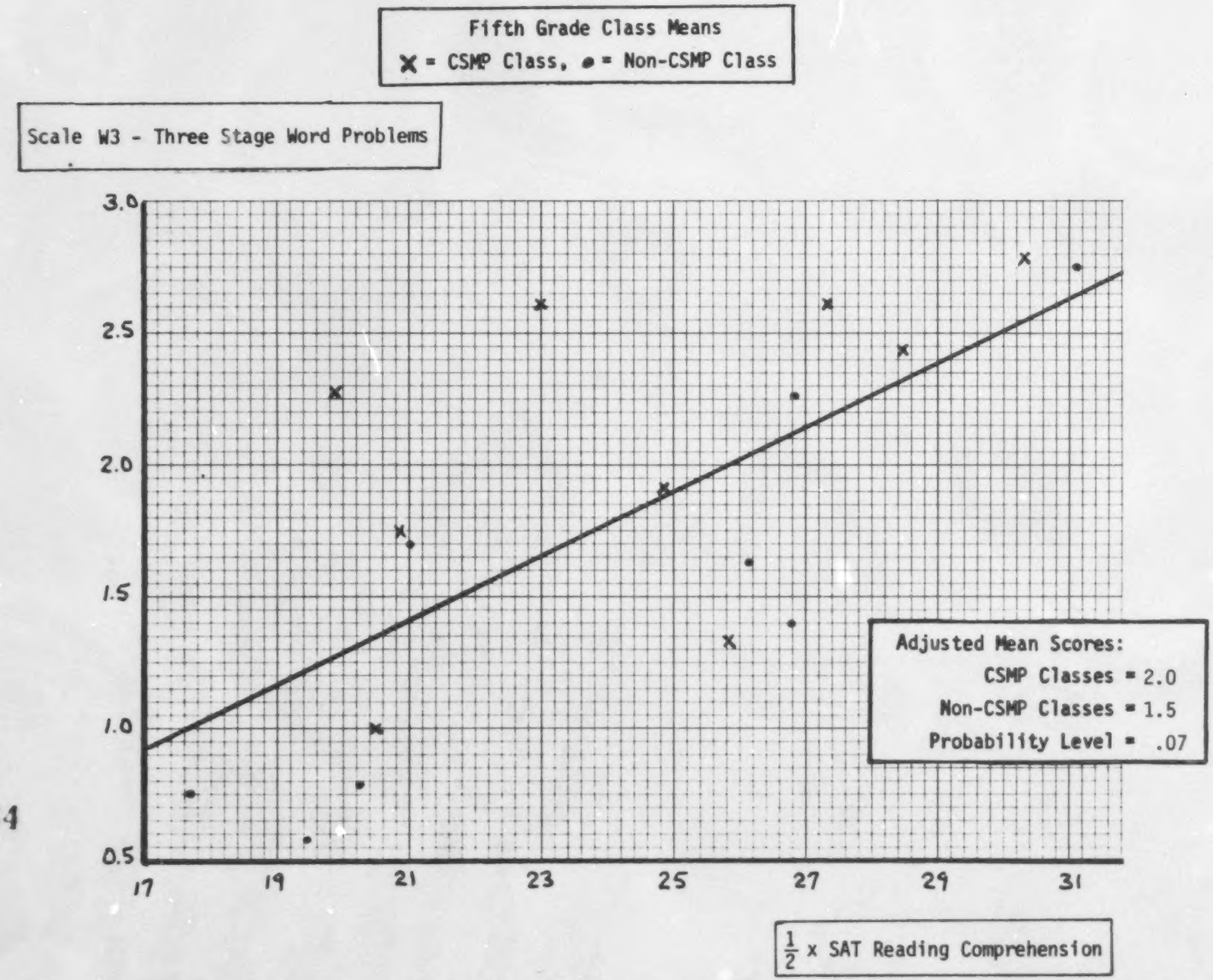
$\frac{1}{2}$ x SAT Reading Comprehension

W3 - Three Stage Word Problems

Test Items	Percent Correct CSMP	Percent Correct Non-CSMP
1. Shirts cost $10 each and ties cost $5 each. Altogether Joe spent $35 for shirts and ties. He bought 2 shirts. How many ties did he buy?_____	77	64
2. Joe puts boxes into 3 piles. Each box is 6 inches high. Each pile is 5 feet high. How many boxes does he need? _____	23	22
3. Bill loads 6 boxes in 2 hours. John loads 4 boxes in 2 hours. Together, how many boxes do they load in 5 hours? _____	24	9
4. Mary has 4 more marbles than Pete. Pete has 2 more marbles than Lisa. Lisa has 3 more marbles than Ed. If Mary has 20 marbles, how many does Ed have? _____	47	30
5. Monday, Tom ran 13 miles. Tuesday, he ran 8 miles. Wednesday, he ran some more. His average for the three days was 10 miles. How many miles did he run on Wednesday? _____	32	20
Number of Students	99	81
KR20 Reliability/Homogeneity	.59	.63
Mean Reading Score	20.1	19.5
Correlation of Scale and Reading Scores	.41	.57

Means by Ability Level

	1	2	3	4	Combined
CSMP ($\bar{X}_1$)	1.5	1.7	2.3	3.0	2.2
Non-CSMP ($\bar{X}_2$)	0.7	1.2	1.2	2.5	1.4
$(\bar{X}_1-\bar{X}_2)/sd$	1.9	1.6	3.3	0.5	3.4


Fifth Grade Class Means
× = CSMP Class, • = Non-CSMP Class
Scale W3 - Three Stage Word Problems
3.0
2.5
2.0
1.5
1.0
0.5
17
19
21
23
25
27
29
31
Adjusted Mean Scores:
CSMP Classes = 2.0
Non-CSMP Classes = 1.5
Probability Level = .07
½ x SAT Reading Comprehension

Appendix B

This appendix contains slightly reduced reproductions of the student pages for three of the MANS scales: those three that could not be fairly represented in the left hand pages of Appendix A. Along with the page reproductions are some commentaries to give a better idea of the actual testing situation.

Contents: R6 - Use of Data

P1 - Spinners

G1 - Geometric Congruencies

(R6) Use of Data (7 items)

Abstract: Students are instructed as to how the "Cost of Pipe" table is read (see page reproduced below). The first two sample items are done as a class with much tester help followed by the students doing the practice item. These first three items are all possible by just reading directly from the table. Then students are shown that all remaining items can't be taken directly from the table but must be estimated (by interpolation or extrapolation). The sample item is explained by the tester. Students then do the last 7 items on their own.

Cost of Pipe

Width \ Length	100'	300'	600'	1,000'
4"	$50	$150	$300	$500
8"	$70	$210	$420	$700
12"	$90	$270	$540	$900
16"	$110	$330	$660	$1100

HOW MUCH DOES IT COST TO BUY PIPE WHICH IS:

Sample: 8" by 600' $_____

Sample: 4" by 300' $_____

Practice: 16" by 100' $_____

HOW MUCH DOES IT COST TO BUY PIPE WHICH IS:

Sample: 10" by 100' $_____

4" by 200' $_____

8" by 100' $_____

10" by 1000' $_____

4" by 500' $_____

4" by 800' $_____

20" by 1000' $_____

6" by 200' $_____

(P1) Spinners

Abstract: On page one, one spinner is shown, and one imagines it being spun 100 times. Then a series of five questions are asked: "about how many times would we get a score of _____?"

On page two, two spinners are shown, with a similar circumstance imagined and a similar 5 questions asked except that the score of the two spinners are to be added. Maximum score is 19 due to credit for relatively correct answers.

SUPPOSE WE PLAY THE GAME 100 TIMES

About how many times would we get a score of 1? ___

About how many times would we get a score of 2? ___

About how many times would we get a score of 3? ___

About how many times would we get a score of 4? ___

About how many times would we get a score of less than 4? ___

SCORES YOU COULD GET 25, 17, 35, 15, 27, 37

SUPPOSE WE PLAY THE GAME 100 TIMES

About how many times would we get a score of 15? ___

About how many times would we get a score of 17? ___

About how many times would we get a score of 25? ___

About how many times would we get a score of less than 20? ___

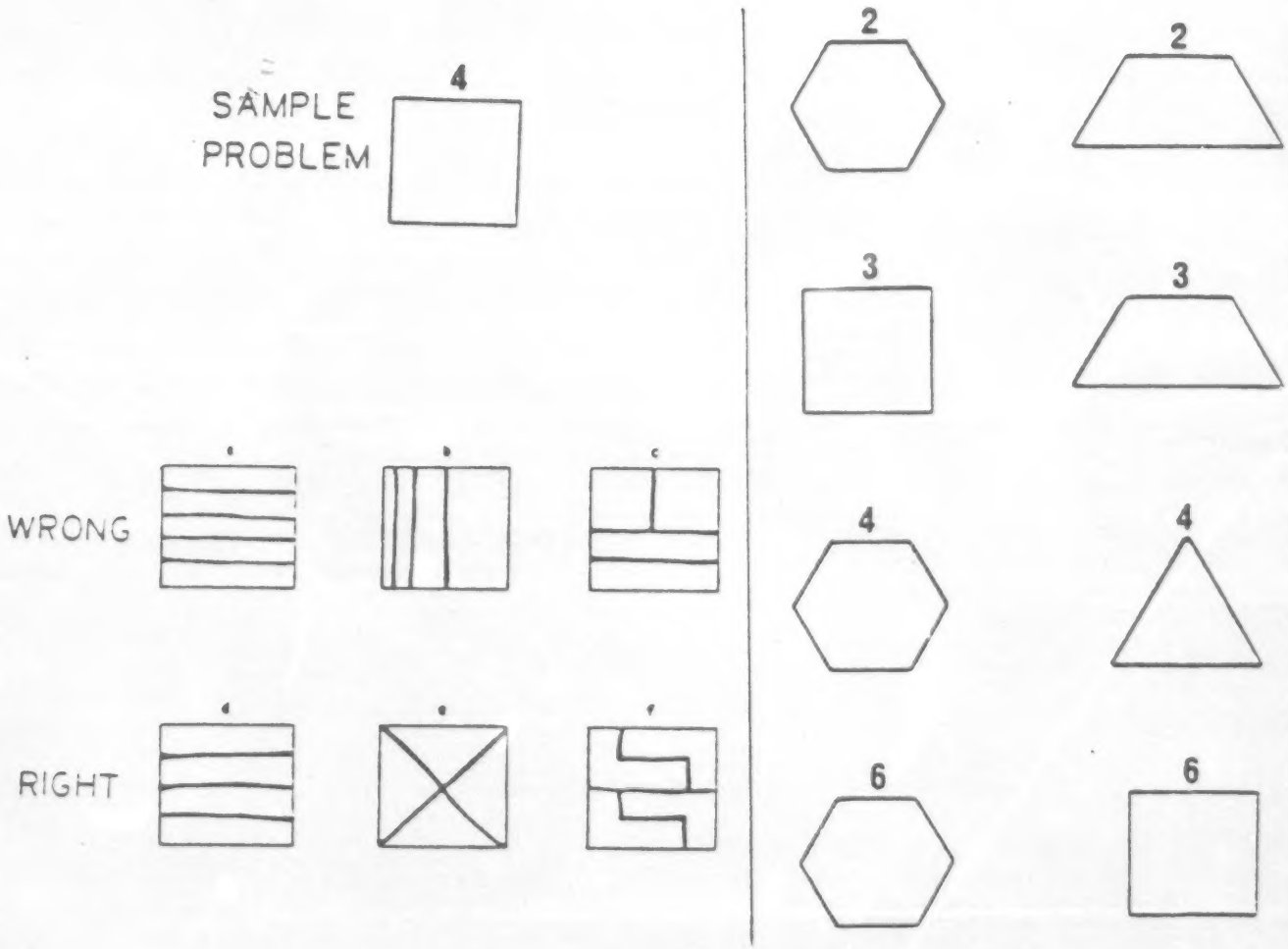
About how many times would we get a score of less than 30? ___

Directions: We play a game by spinning the spinner. On the second page, we spin both spinners at the same time, and the score is the total from the two spinners.

General Goal: Use elementary ideas in probability to estimate how frequently various results will occur.

(G1) Geometric Congruencies

Abstract: After examining 3 correct and 3 incorrect solutions to a sample problem, given a regular geometric shape and a number of parts, the shape must be divided into that many congruent parts, 8 items.



Directions: Draw lines to divide each figure into the indicated number of congruent parts. The first page was done as an illustration.

General Goal: Visualize congruent parts of common geometric shapes.